STP – Spanning Tree Protocol

Cabrillo College

CIS 187 Multilayer Switched Networks

CCNP 3 version 4

Rick Graziani

Fall 2006

Spanning Tree Protocol (STP)

Cabrillo College

- STP often accounts for more than 50 % of the configuration, troubleshooting, and maintenance headaches in real-world campus networks (especially if they are poorly designed).
- Complex protocol that is generally poorly understood.
- Radia Perlman – Developer of STP

Configuring STP

- By default, STP is enabled for every port on the switch.
- If for some reason STP has been disabled, you can reenable it.
- To re-enable STP, use the

Switch(config)#spanning-tree vlan *vlan-id*

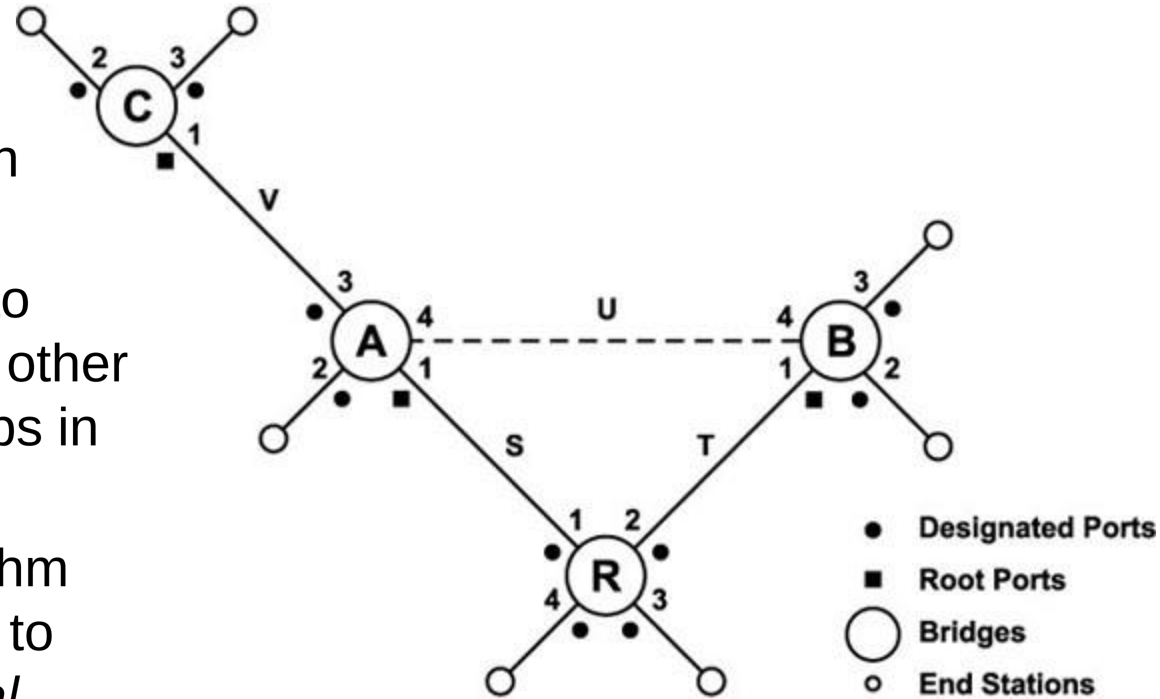
- To disable STP, on a per-VLAN basis:

Switch(config)#no spanning-tree vlan *vlan-id*

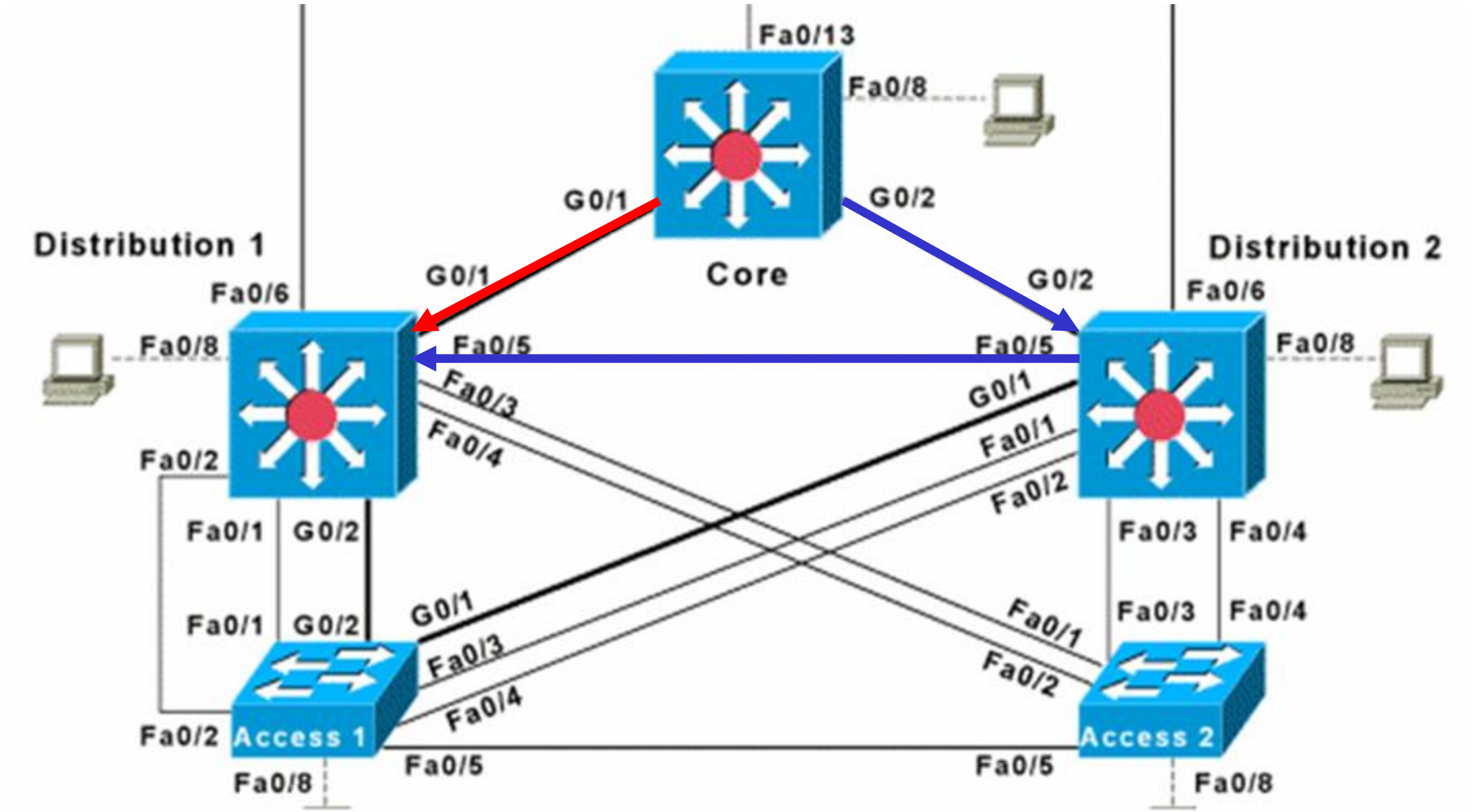
Spanning Tree Protocol (STP)

Cabrillo College

- STP is a loop-prevention protocol
- STP allows L2 devices to communicate with each other to discover physical loops in the network.
- STP specifies an algorithm that L2 devices can use to create a *loop-free logical* topology.
- STP creates a tree structure of loop-free leaves and branches that spans the entire Layer 2 network.



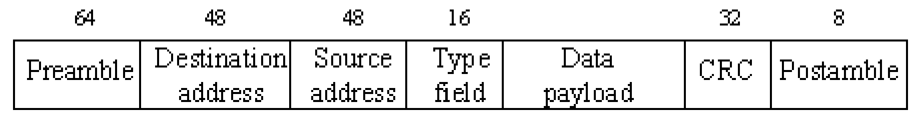
Redundancy Creates Loops



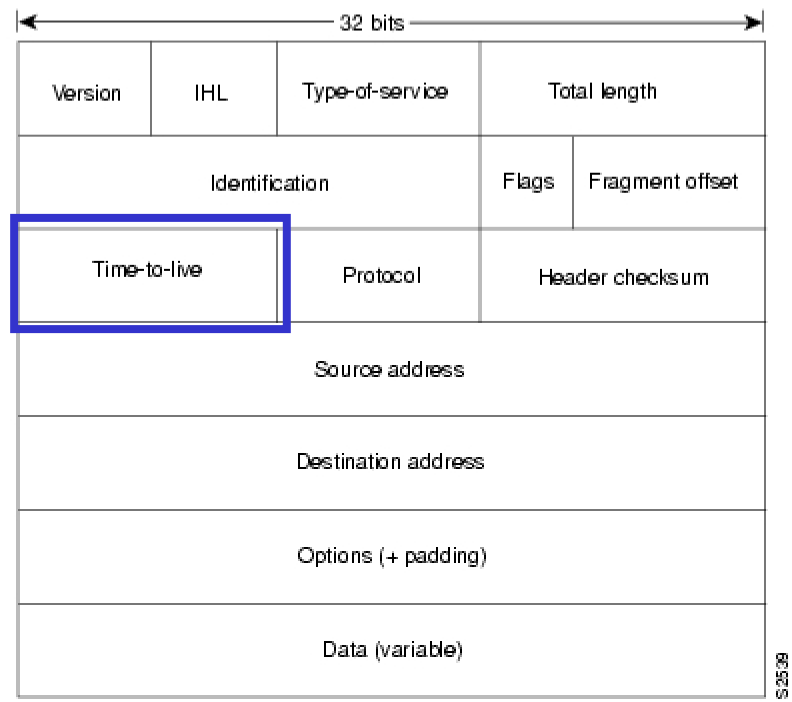
L2 Loops

- Broadcasts and Layer 2 loops can be a dangerous combination.
- Ethernet frames have no TTL field
- After an Ethernet frame starts to loop, it will probably continue until someone shuts off one of the switches or breaks a link.

Ethernet Frame Format

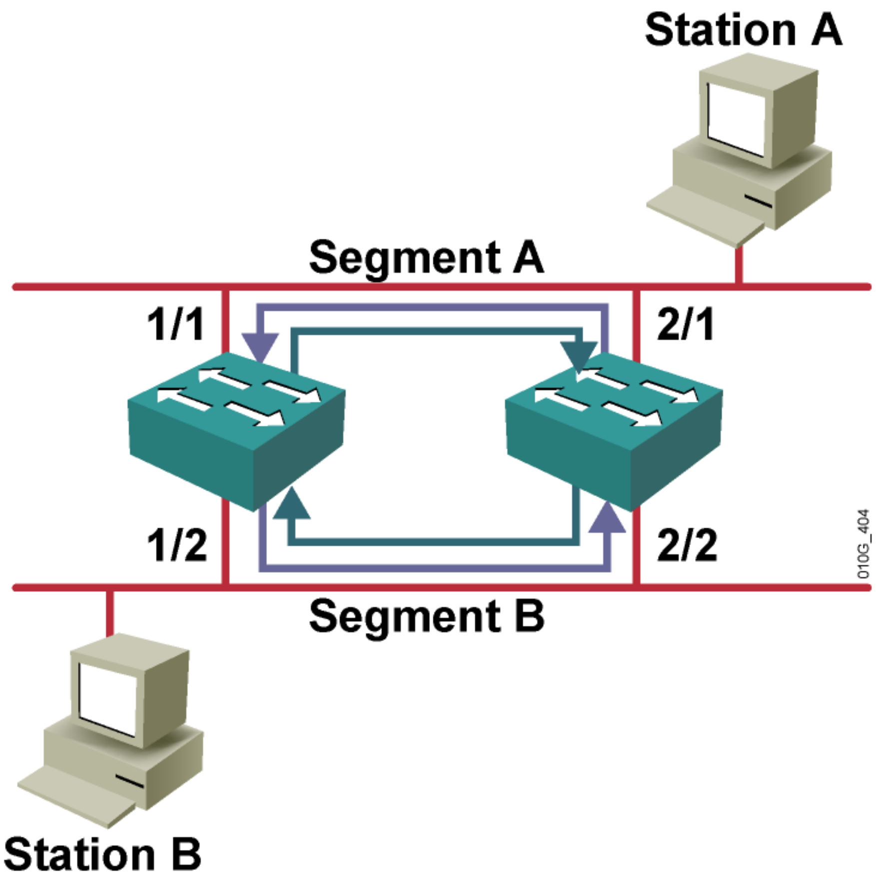


IP Packet



L2 Loops - Flooded unicast frames

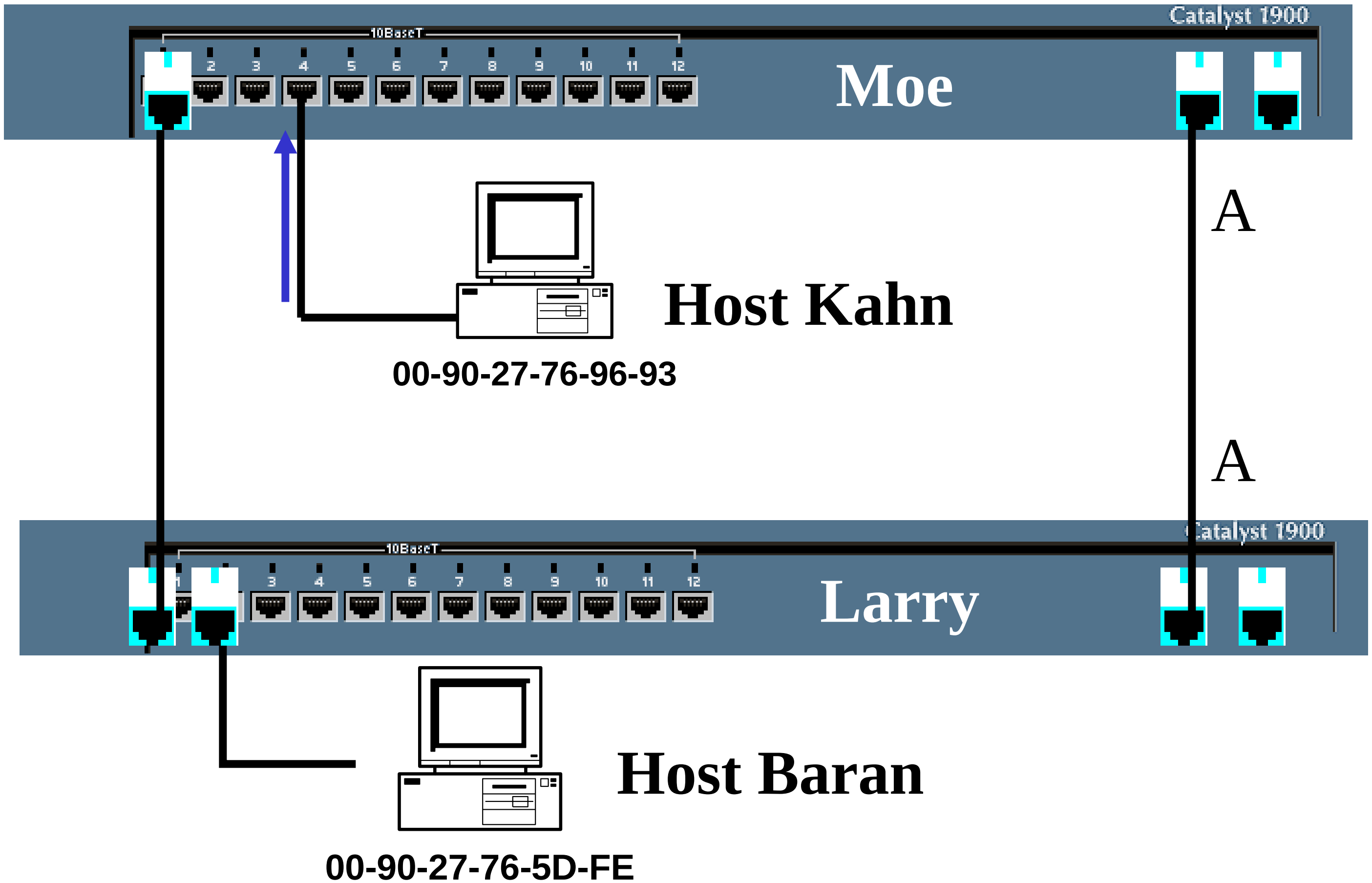
- Bridge loops can occur any time there is a redundant path or loop in the bridge network.
- The switches will flip flop the bridging table entry for Station A (creating extremely high CPU utilization).



Unknown Unicast

Switch Moe learns Kahns' MAC address.

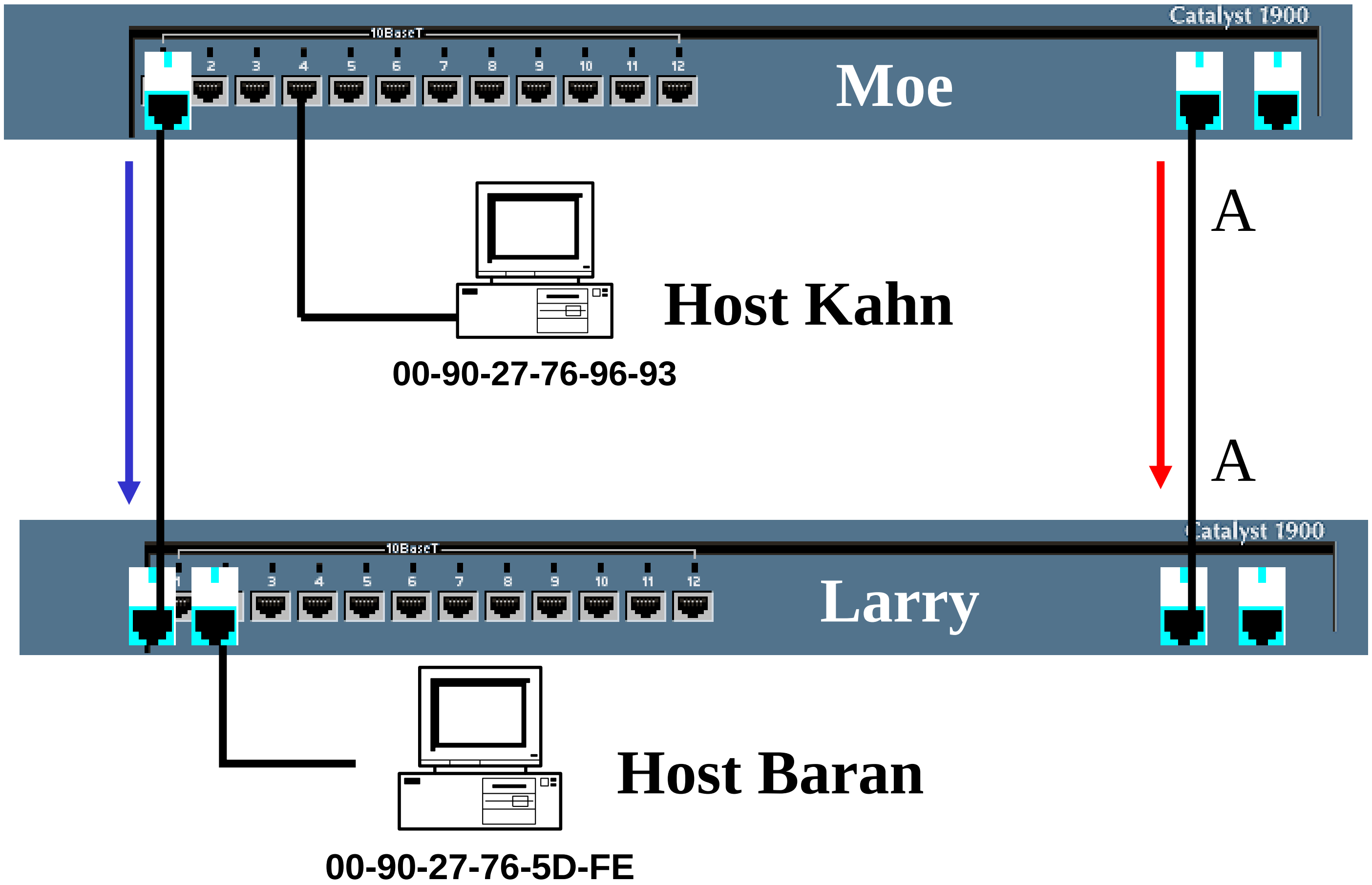
```
SAT (Source Address Table)
Port 4:  00-90-27-76-96-93
```



Unknown Unicast

- Destination MAC is an unknown unicast, so Moe floods it out all ports.

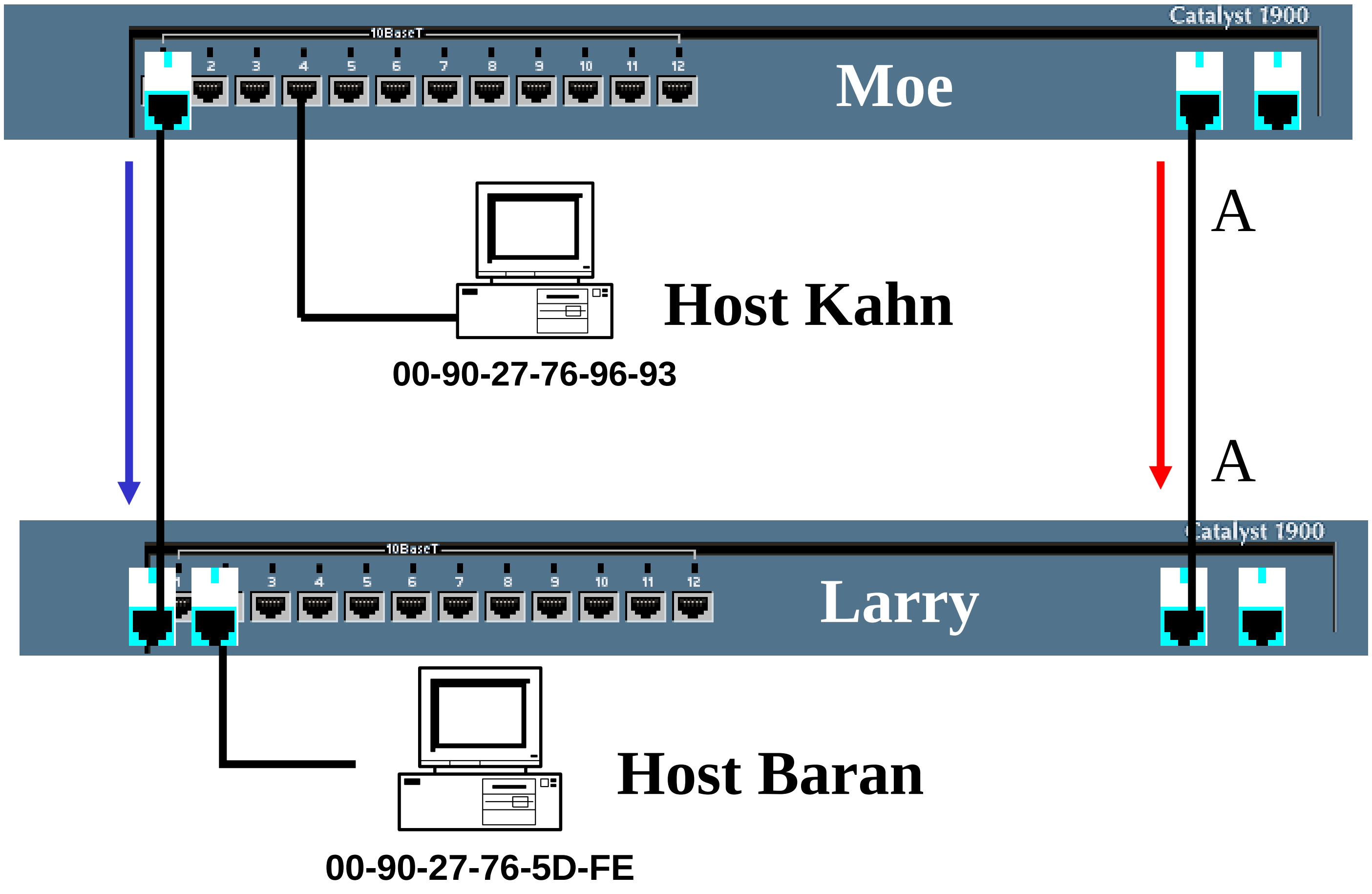
```
SAT (Source Address Table)

Port 4:   00-90-27-76-96-93
```



Unknown Unicast

- Destination MAC is an unknown unicast, so Moe floods it out all ports.

```
SAT (Source Address Table)
Port 4:   00-90-27-76-96-93
```

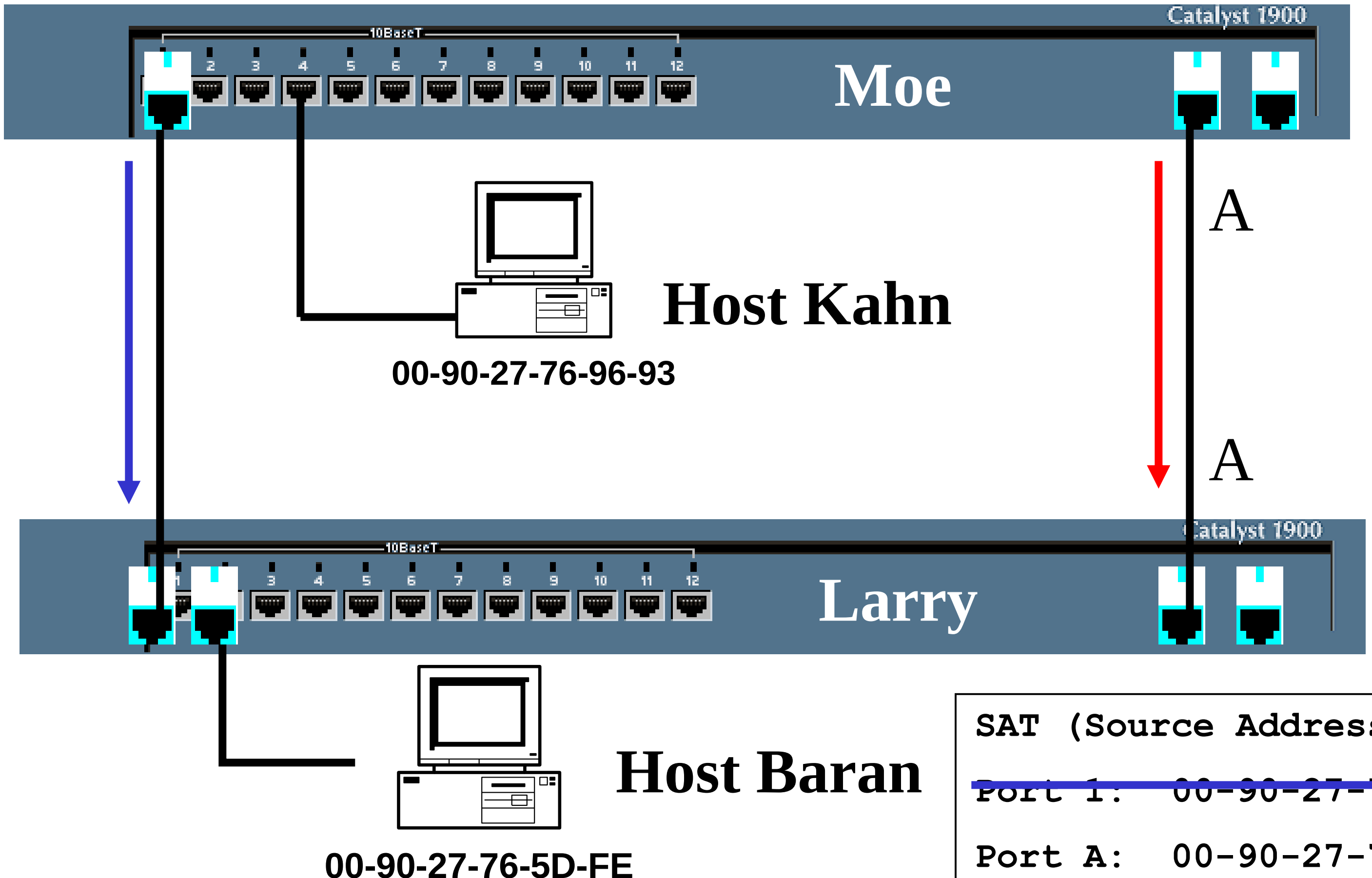


Unknown Unicast

Cabrillo College

- Switch Larry records the Source MAC of the frame twice with the last one being the most recent.

SAT (Source Address Table)	
Port 4:	00-90-27-76-96-93



SAT (Source Address Table)	
~~Port 1:~~	~~00-90-27-76-96-93~~
Port A:	00-90-27-76-96-93

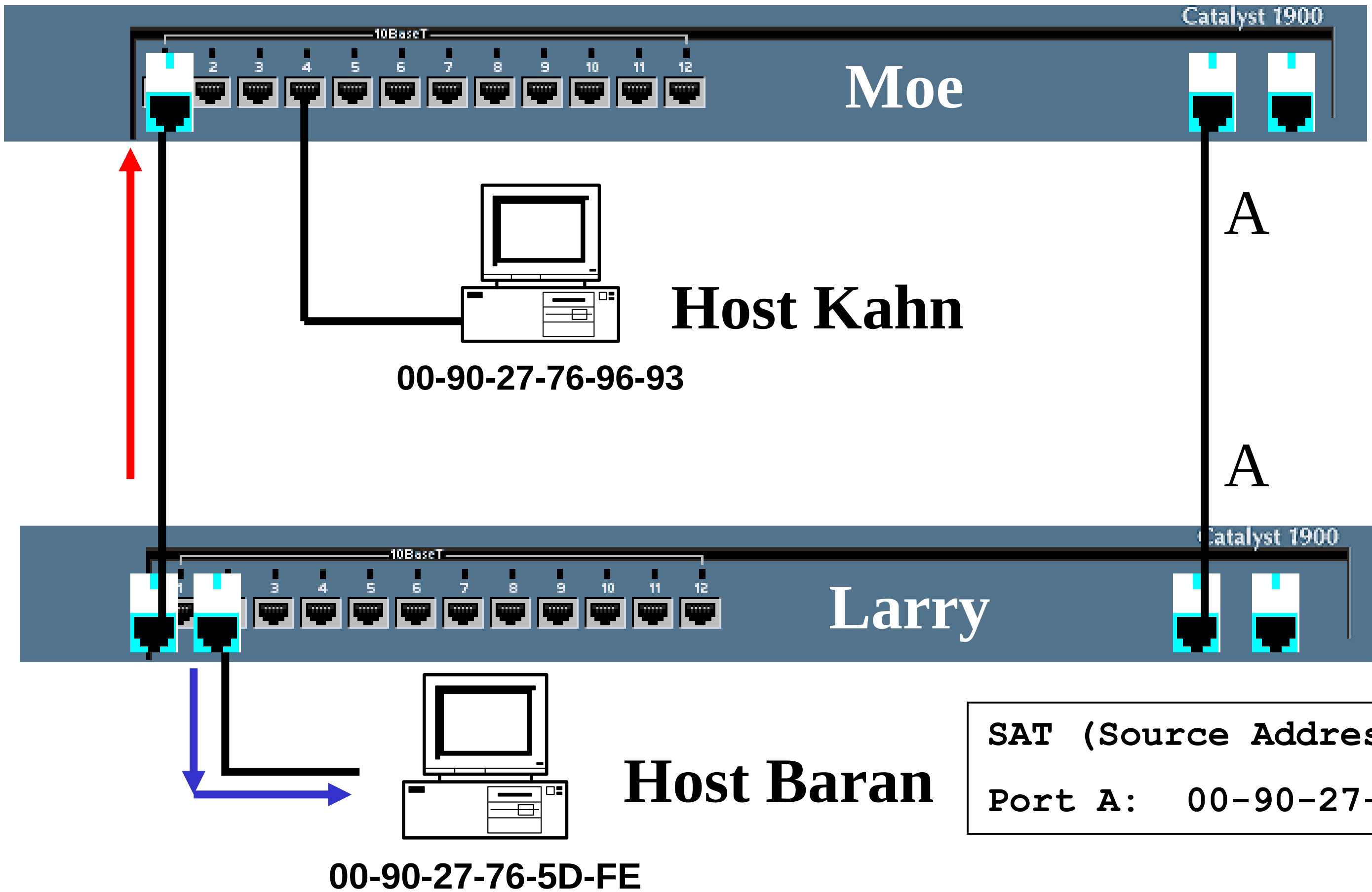
Unknown Unicast

Cabrillo College

- Switch Larry floods the unknown unicast out all ports, except the incoming port.

SAT (Source Address Table)

Port 1: 00-90-27-76-96-93



SAT (Source Address Table)

Port A: 00-90-27-76-96-93

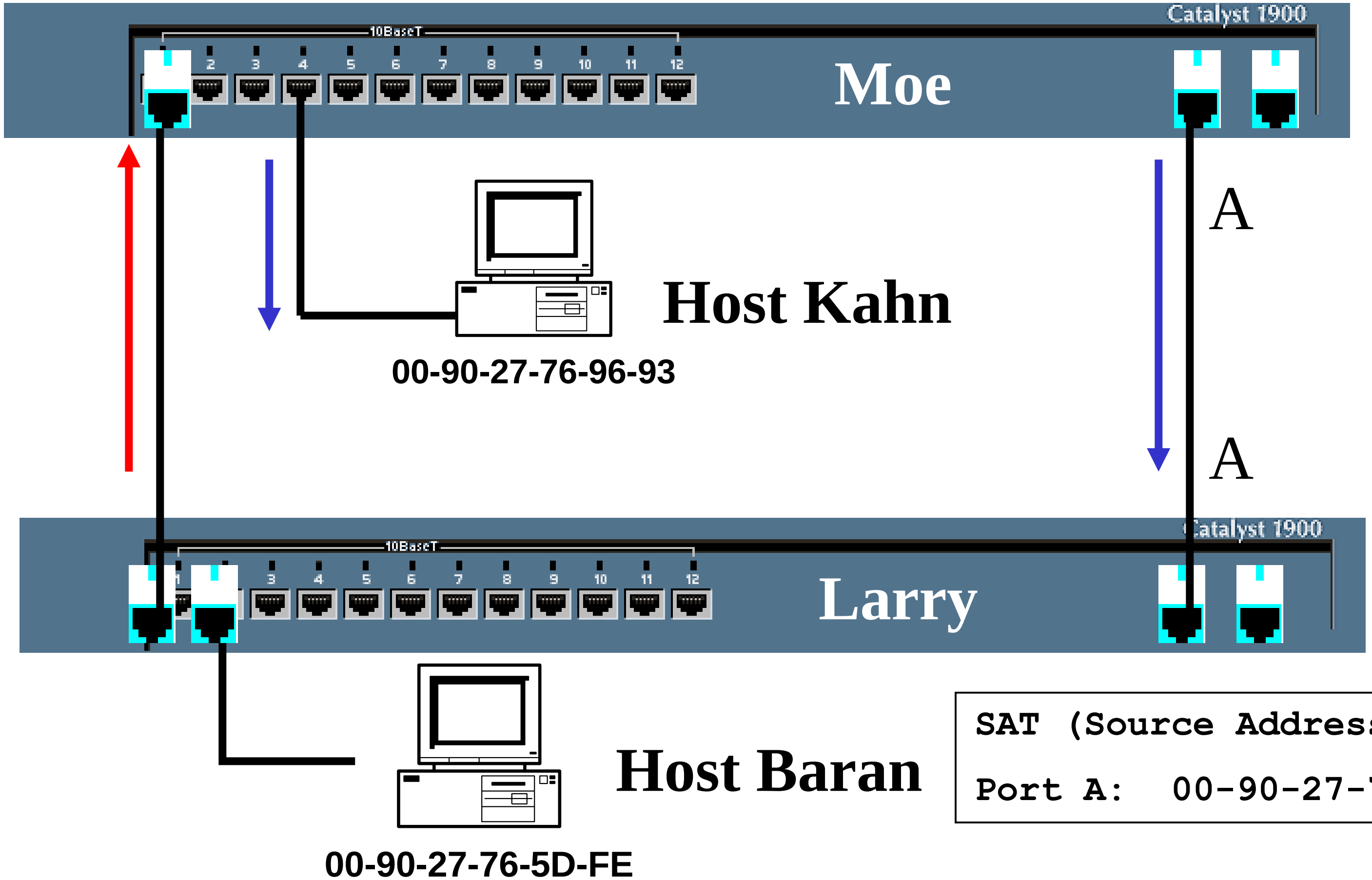
Unknown Unicast

- Switch Moe receives the frame, changes the MAC address table with newer information and floods the unknown unicast out all ports.

SAT (Source Address Table)

~~Port 4: 00-90-27-76-96-93~~

Port 1: 00-90-27-76-96-93



SAT (Source Address Table)

Port A: 00-90-27-76-96-93

Unknown Unicast

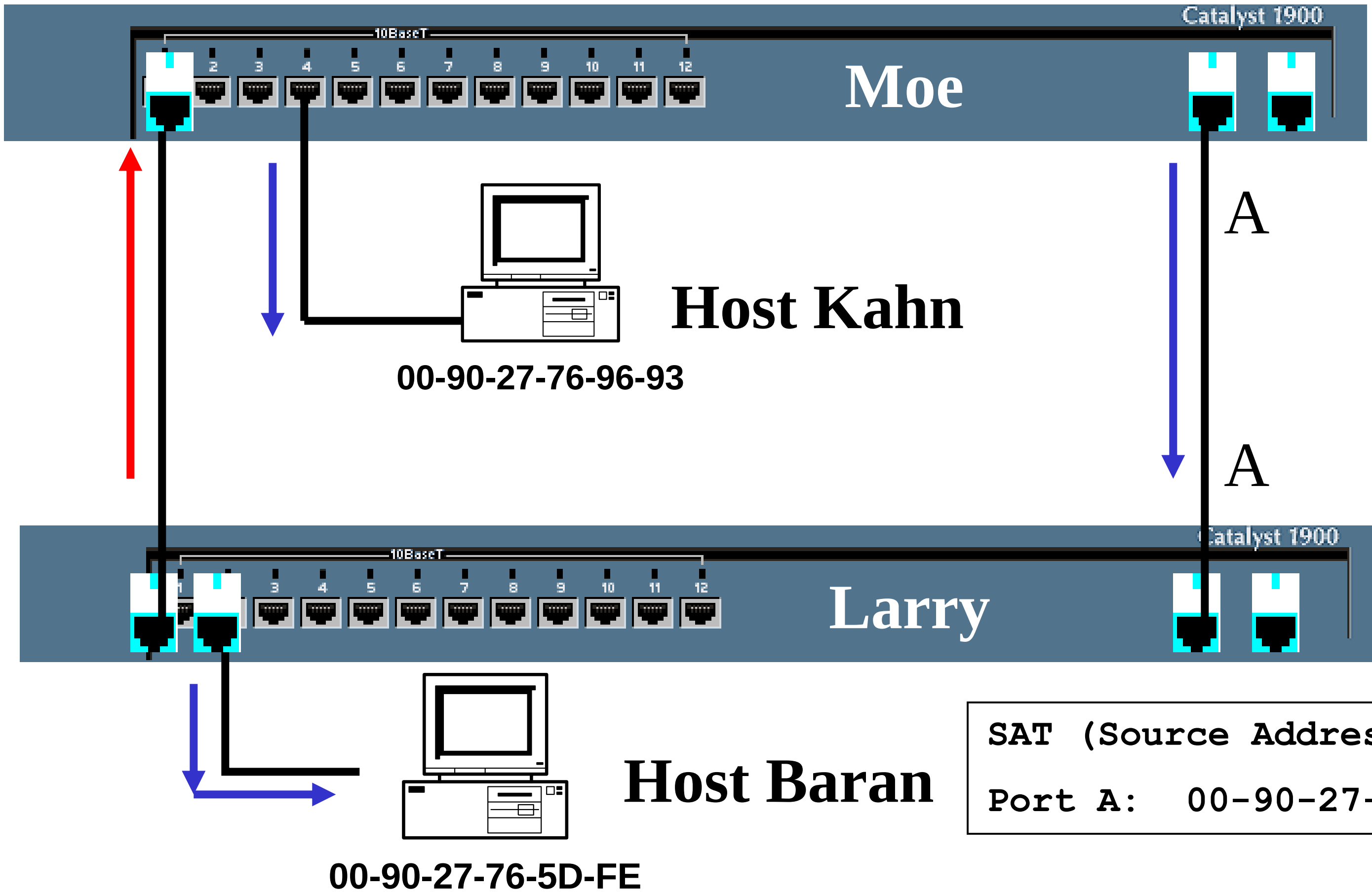
Cabrillo College

- And the cycle continues!

SAT (Source Address Table)

~~Port 4: 00-90-27-76-96-93~~

Port 1: 00-90-27-76-96-93



SAT (Source Address Table)

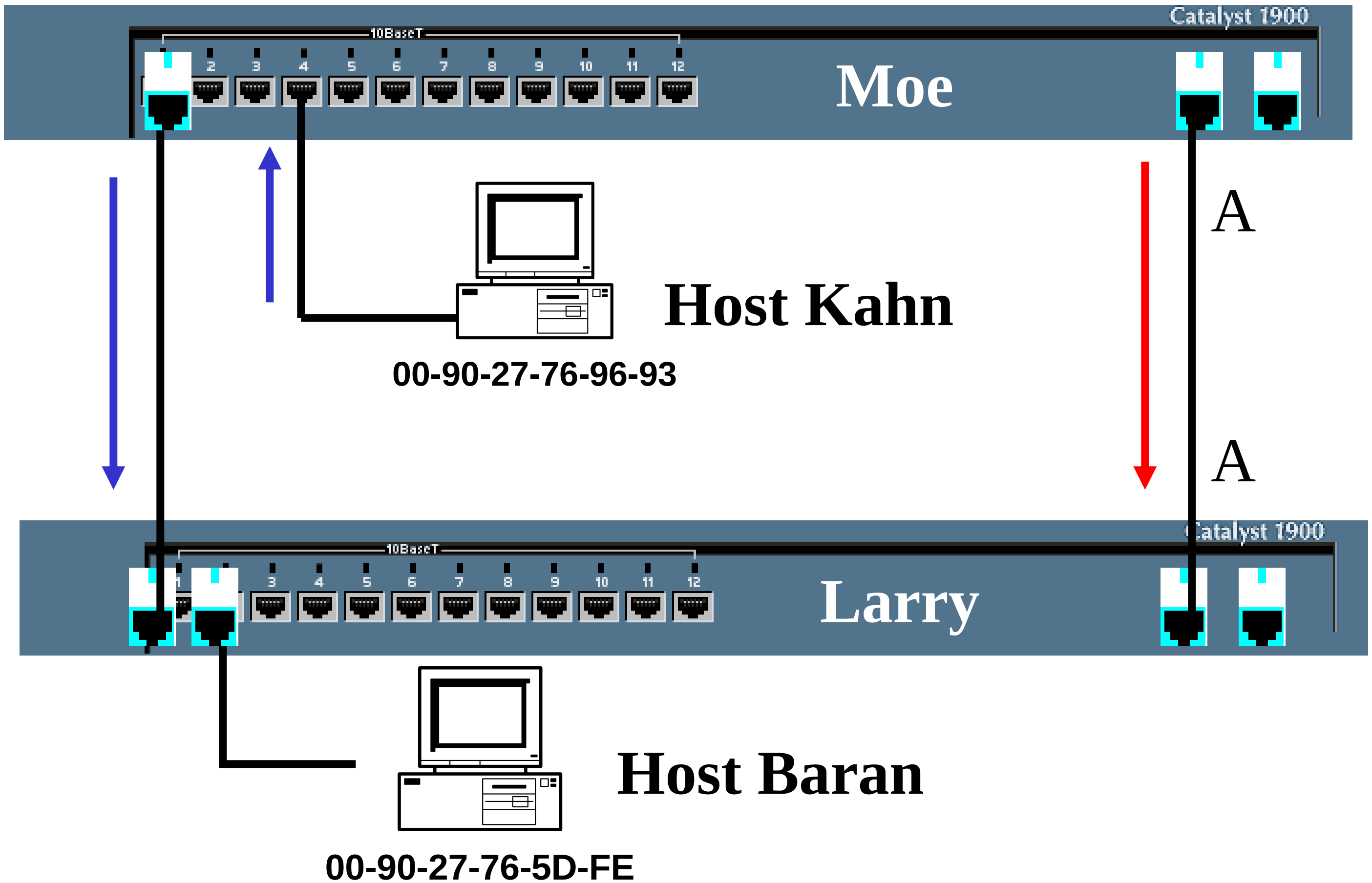
Port A: 00-90-27-76-96-93

Layer 2 Broadcast

Cabrillo College

- Host Kahn sends an ARP Request, a Layer 2 broadcast

SAT (Source Address Table)	
Port 1:	00-90-27-76-96-93

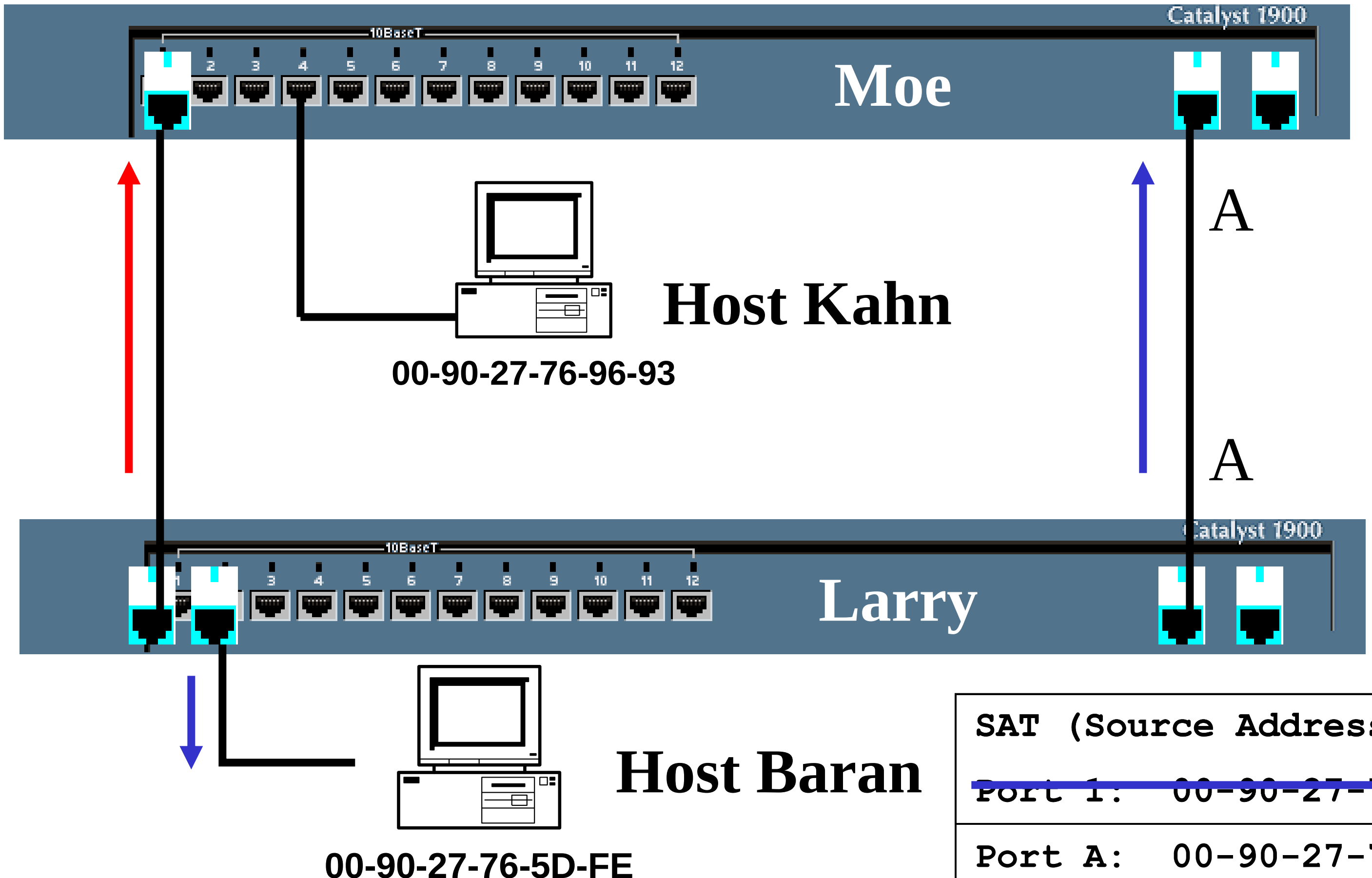


Layer 2 Broadcast

Cabrillo College

- Switch Moe floods the frame.
- Switch Larry floods the frames.
- Switches continue to flood duplicate frames.
- Switches constantly modifying MAC Address Tables

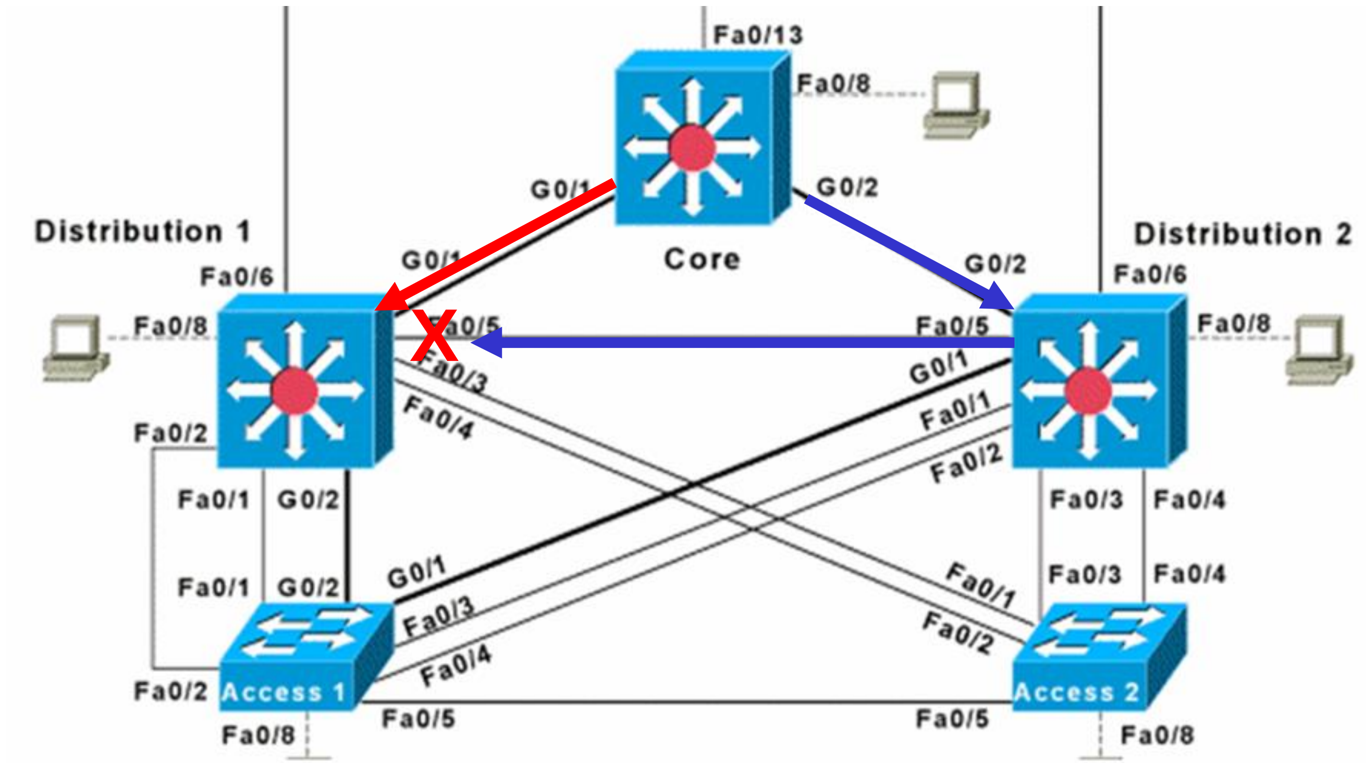
SAT (Source Address Table)	
Port 1:	00-90-27-76-96-93



SAT (Source Address Table)	
~~Port 1:~~	~~00-90-27-76-96-93~~
Port A:	00-90-27-76-96-93

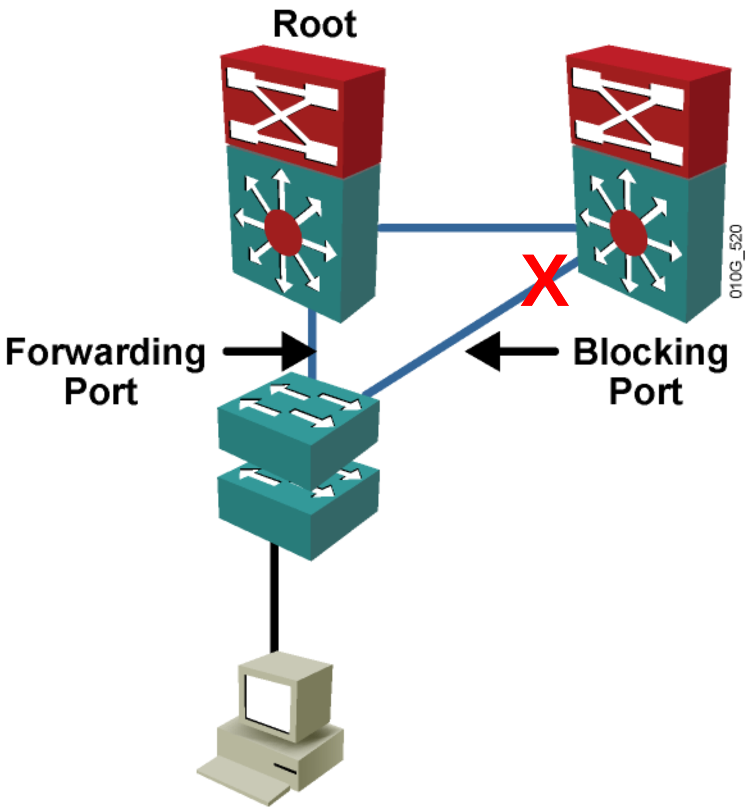
STP Prevents Loops

- The purpose of STP is to avoid and eliminate loops in the network by negotiating a loop-free path through a root bridge.
- STP determines where the are loops and blocks links that are redundant.
- Ensures that there will be only one active path to every destination.



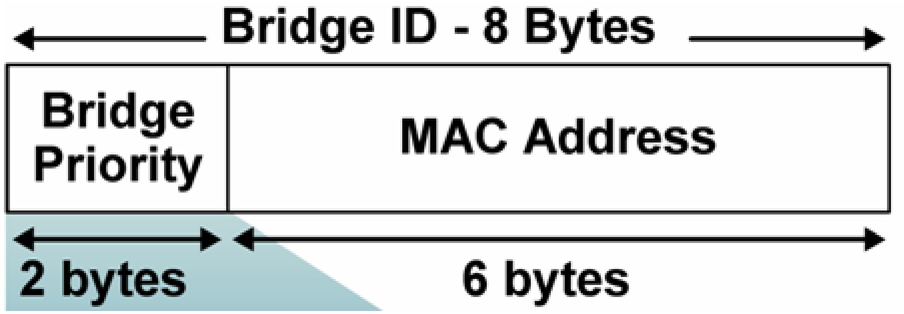
Spanning Tree Algorithm

- STP executes an algorithm called STA.
- STA chooses a reference point, called a root bridge, and then determines the available paths to that reference point.
- If more than two paths exists, STA picks the best path and blocks the rest



Two-key STP Concepts

- STP calculations make extensive use of two key concepts in creating a loop-free topology:
  - Bridge ID
  - Path Cost



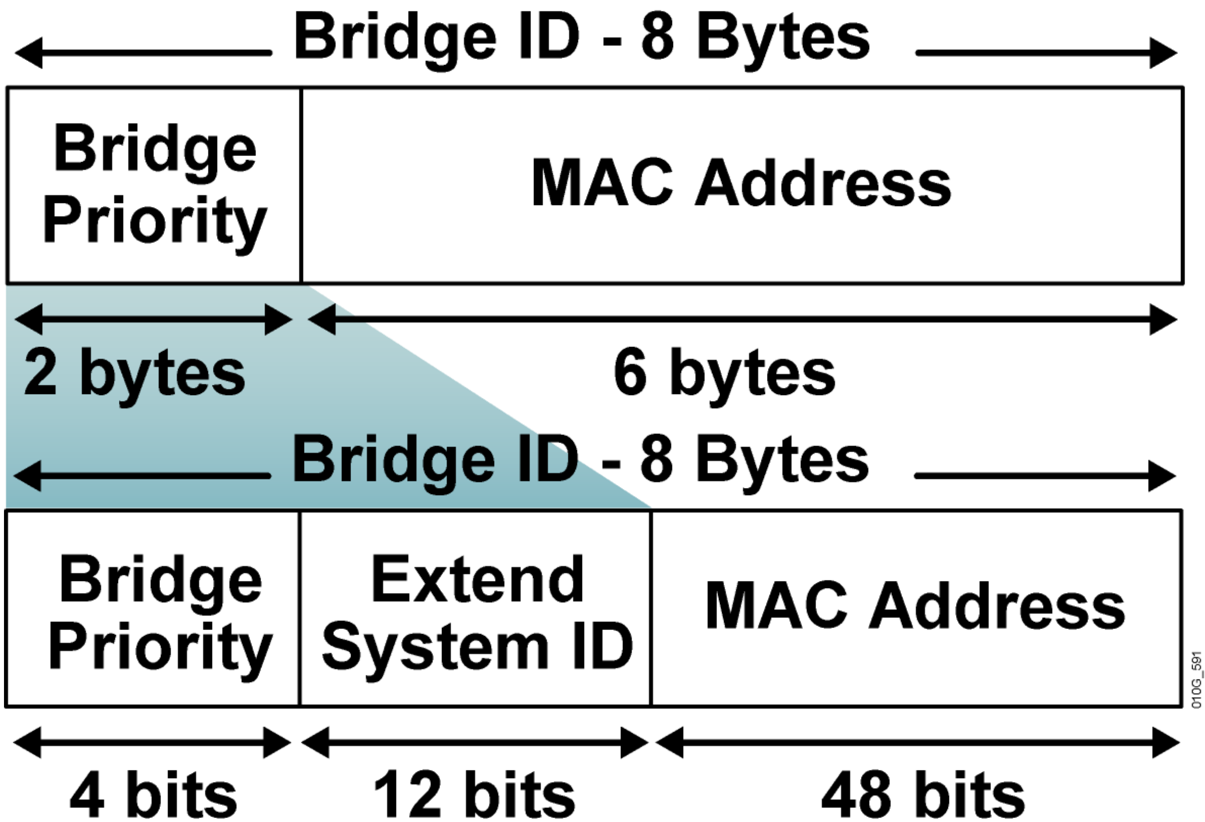
Link Speed	Cost (Revised IEEE Spec)	Cost (Previous IEEE Spec)
10 Gbps	2	1
1 Gbps	4	1
100 Mbps	19	10
10 Mbps	100	100

Bridge ID (BID)

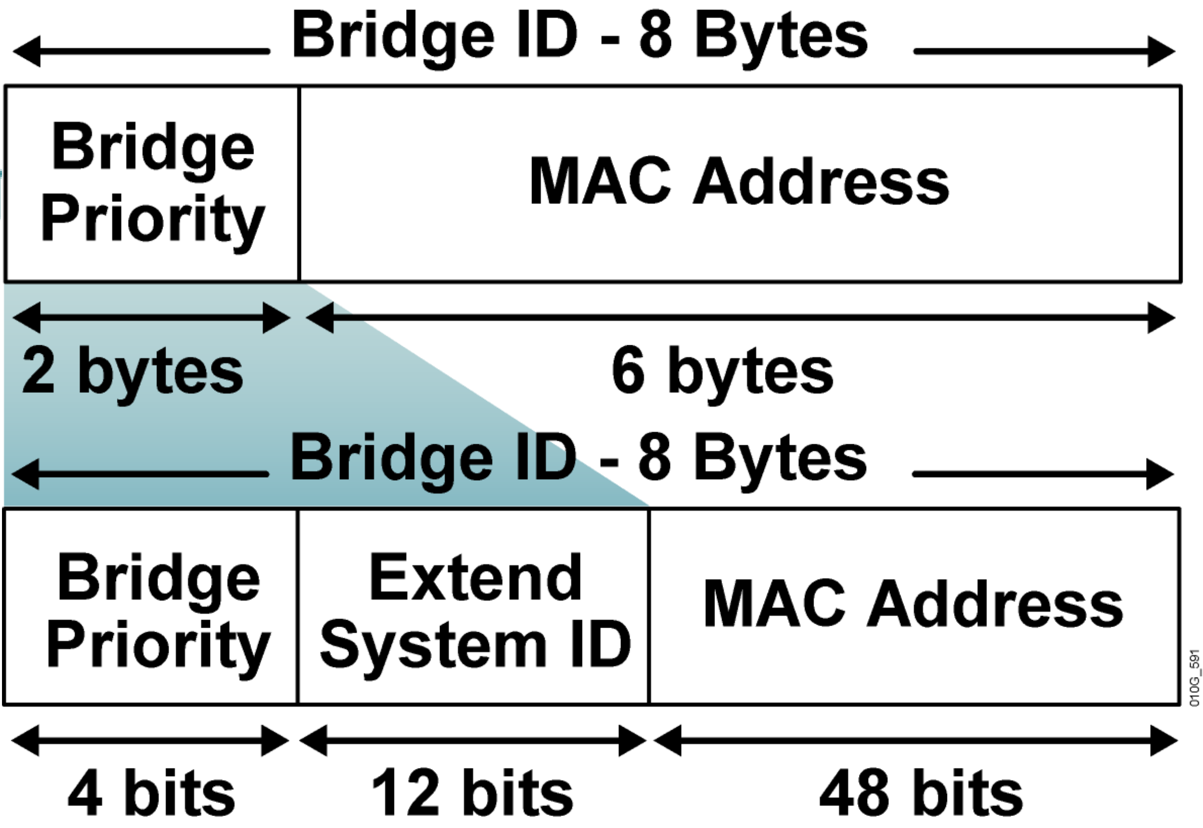
- Bridge ID (BID) is used to identify each bridge/switch.
- The BID is used in determining the center of the network, in respect to STP, known as the root bridge.

Bridge ID Without the Extended System ID

Bridge ID with the Extended System ID

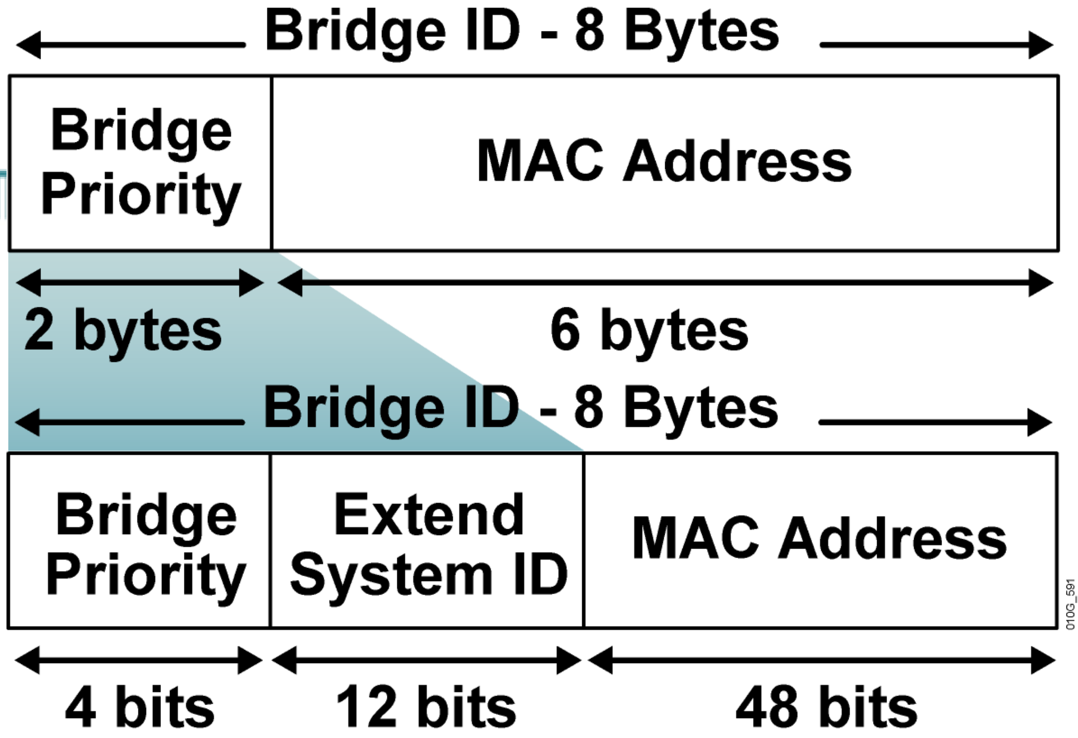


Bridge ID (BID)



- Consists of two components:
  - **A 2-byte Bridge Priority**: Cisco switch defaults to 32,768 or 0x8000.
  - **A 6-byte MAC address**
- **Bridge Priority** is usually expressed in **decimal format** and the **MAC address** in the BID is usually expressed in **hexadecimal format**.

Bridge ID (BID)



- Spanning tree operation requires that each switch have a unique BID.
- In the **original 802.1D** standard, the BID was composed of the **Priority** Field and the **MAC address** of the switch, and all VLANs were represented by a CST.
- Because **PVST** requires that a separate instance of spanning tree run for each VLAN, the **BID field is required to carry VLAN ID (VID)** information.
- This is accomplished by **reusing a portion of the Priority** field as the extended **system ID to carry a VID**.

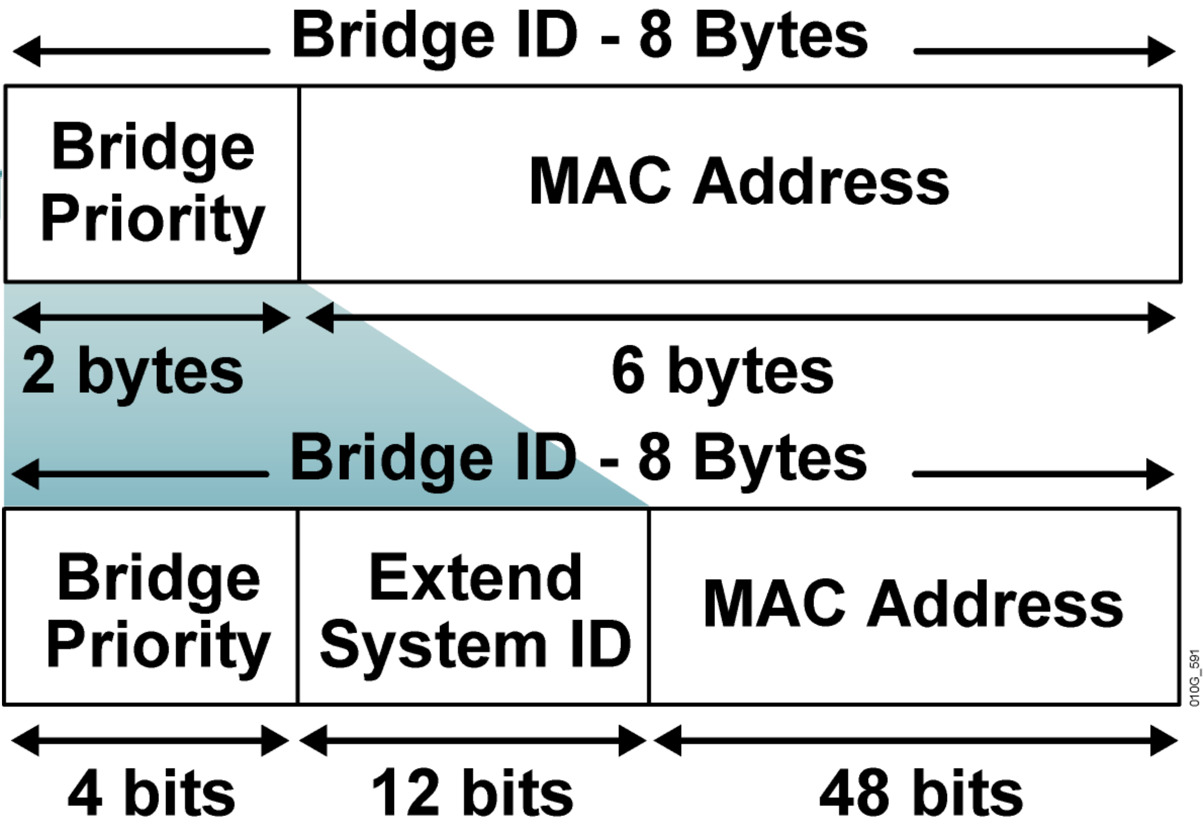
Priority = Priority *(Default 32,768)* + VLAN

Cabrillo College

```
Access2#show spanning-tree
VLAN0001
  Spanning tree enabled protocol ieee
  Root ID    Priority    24577
             Address     000f.2490.1380
             Cost        23
             Port        1 (FastEthernet0/1)
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
  Bridge ID  Priority    32769  (priority 32768 sys-id-ext 1)
             Address     0009.7c0b.e7c0
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
             Aging Time 300
<text omitted>

VLAN0010
  Spanning tree enabled protocol ieee
  Root ID    Priority    4106
             Address     000b.fd13.9080
             Cost        19
             Port        1 (FastEthernet0/1)
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
  Bridge ID  Priority    32778  (priority 32768 sys-id-ext 10)
             Address     0009.7c0b.e7c0
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
             Aging Time 300
```

PVST coming later…

Bridge ID (BID)



- Used to elect a root bridge (coming)
- **Lowest** Bridge ID is the root.
- If all devices have the same priority, the bridge with the lowest MAC address becomes the root bridge. (Yikes)
- **Note**: For simplicity, in our topologies we will use Bridge Priorities without the Extended System ID.

Path Cost – Original Spec (Linear)

Cabrillo College

Link Speed	Cost (Revised IEEE Spec)	Cost (Previous IEEE Spec)
10 Gbps	2	1
1 Gbps	4	1
100 Mbps	19	10
10 Mbps	100	100

- Bridges use the concept of cost to evaluate how close they are to other bridges.
- This will be used in the STP development of a loop-free topology .
- Originally, 802.1D defined cost as **1 billion/bandwidth** of the link in Mbps.
  - Cost of 10 Mbps link = 100 or 1000/10
  - Cost of 100 Mbps link = 10 or 1000/100
  - Cost of 1 Gbps link = 1 or 1000/1000
- Running out of room for faster switches including 10 Gbps Ethernet

Path Cost – Revised Spec (Non-Linear)

Cabrillo College

Link Speed	Cost (Revised IEEE Spec)	Cost (Previous IEEE Spec)
10 Gbps	2	1
1 Gbps	4	1
100 Mbps	19	10
10 Mbps	100	100

- IEEE modified the most to use a non-linear scale with the new values of:
  - 4 Mbps 250 (cost)
  - 10 Mbps 100 (cost)
  - 16 Mbps 62 (cost)
  - 45 Mbps 39 (cost)
  - 100 Mbps 19 (cost)
  - 155 Mbps 14 (cost)
  - 622 Mbps 6 (cost)
  - 1 Gbps 4 (cost)
  - 10 Gbps 2 (cost)

- You can change the path cost by modifying the cost of a port.
- Exercise caution when you do this!
- BID and Path Cost are used to develop a loop-free topology .
- Coming very soon!

Five-Step STP Decision Sequence

Cabrillo College

- When creating a loop-free topology, STP always uses the same five-step decision sequence:

Five-Step decision Sequence

Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 – Lowest Port Priority

Step 5 - Lowest Port ID

- Bridges use Configuration BPDUs during this four-step process.
- We will assume all BPDUs are configuration BPDUs until otherwise noted.

Five-Step STP Decision Sequence

BPDU key concepts:

- Bridges save a copy of only the best BPDU seen on every port.
- When making this evaluation, it considers all of the BPDUs received on the port, as well as the BPDU that would be sent on that port.
- As every BPDU arrives, it is checked against this five-step sequence to see if it is more attractive (lower in value) than the existing BPDU saved for that port.
- Only the lowest value BPDU is saved.
- Bridges send configuration BPDUs until a more attractive BPDU is received.
- Okay, lets see how this is used...

Elect one Root Bridge

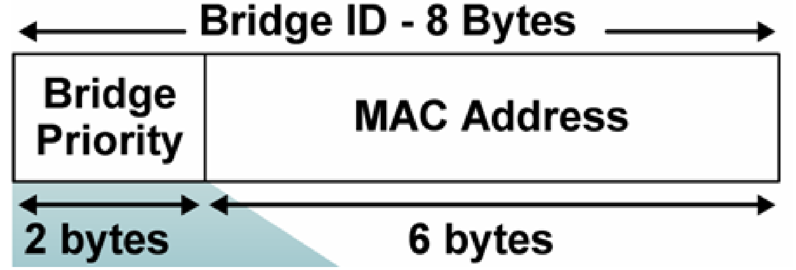
The STP algorithm uses three simple steps to converge on a loop-free topology:

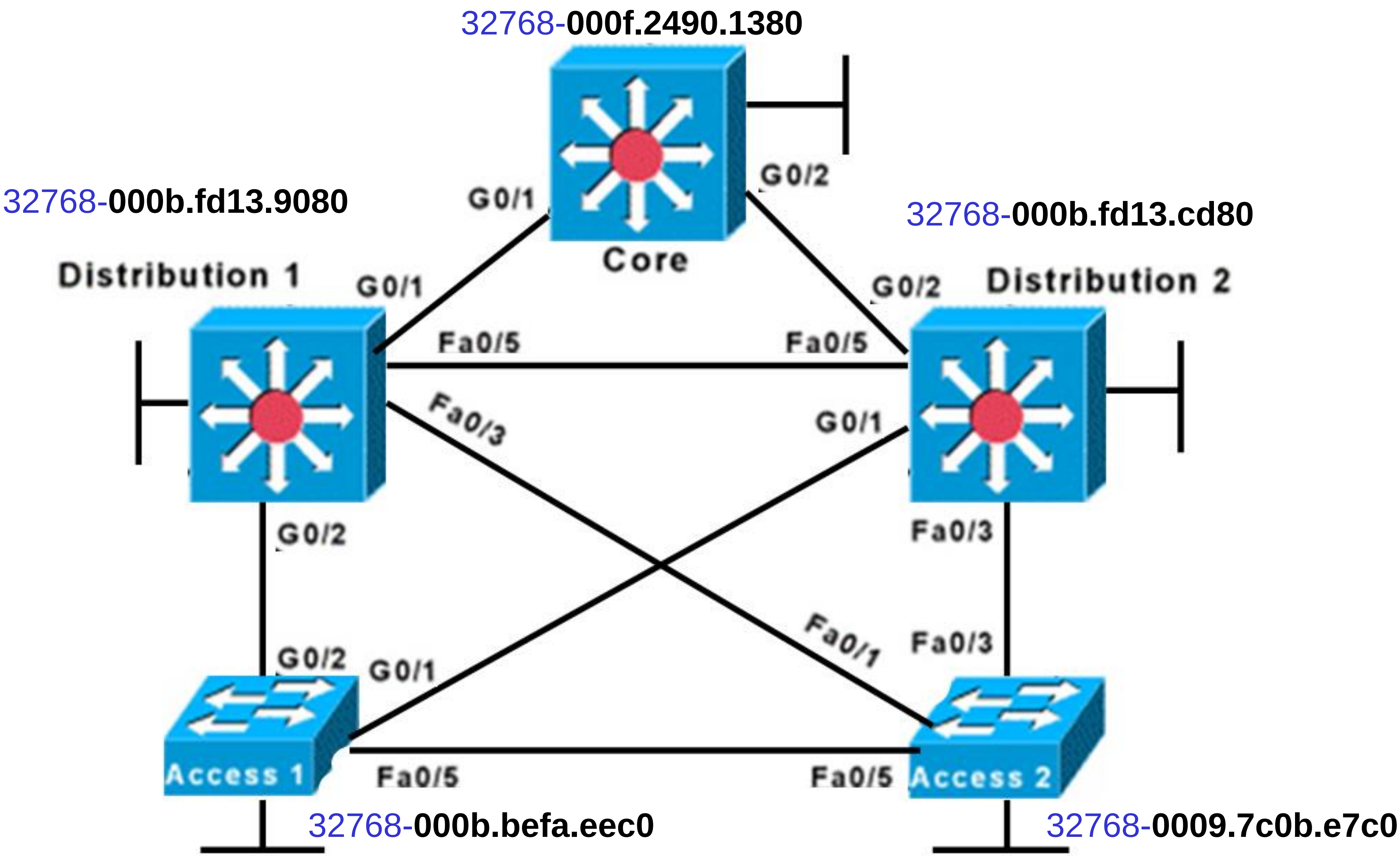
STP Convergence

Step 1 Elect one Root Bridge
Step 2 Elect Root Ports
Step 3 Elect Designated Ports

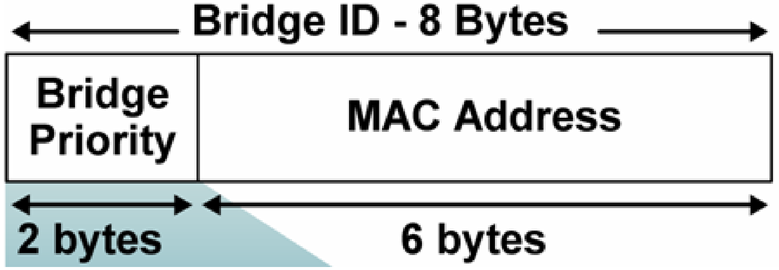
- When the network first starts, all bridges are announcing a chaotic mix of BPDUs.
- All bridges immediately begin applying the five-step sequence decision process.
- Switches need to elect a single Root Bridge.
- Switch with the **lowest BID** wins!
- Note: Many texts refer to the term “highest priority” which is the “lowest” BID value.
- This is known as the **“Root War.”**

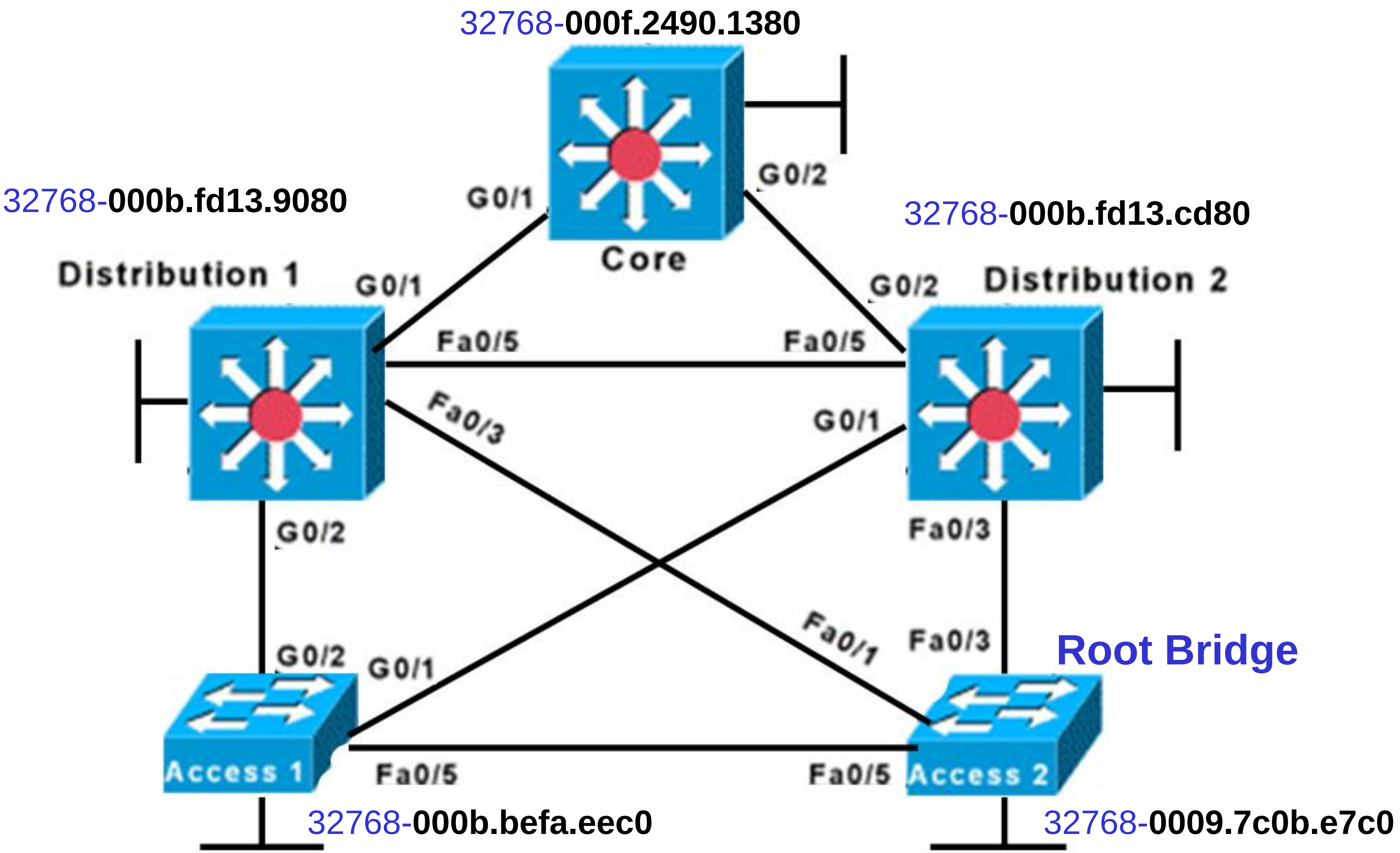
Elect one Root Bridge
Lowest BID wins!





Elect one Root Bridge Lowest BID wins!





Elect one Root Bridge
Lowest BID wins!

Its all done with BPDUs!

Bytes	Field
2	Protocol ID
1	Version
1	Message type
1	Flags
8	Root ID
4	Cost of path
8	Bridge ID
2	Port ID
2	Message age
2	Max age
2	Hello time
2	Forward delay

310P_126

Who is the root bridge?

How far away is the root bridge?

What is the BID of the bridge that sent this BPDU?

What port on the sending bridge did this BPDU come from?

Elect one Root Bridge
Lowest BID wins!


```
BPDU
802.3 Header
  Destination:    01:80:C2:00:00:00  Mcast 802.1d Bridge group
  Source:         00:D0:C0:F5:18:D1
  LLC Length:     38
802.2 Logical Link Control (LLC) Header
  Dest. SAP:      0x42  802.1 Bridge Spanning Tree
  Source SAP:     0x42  802.1 Bridge Spanning Tree
  Command:        0x03  Unnumbered Information
802.1 - Bridge Spanning Tree
  Protocol Identifier:  0
  Protocol Version ID:  0
  Message Type:         0  Configuration Message
  Flags:                %00000000
  Root Priority/ID:     0x8000/ 00:D0:C0:F5:18:C0
  Cost Of Path To Root: 0x00000000  (0)
  Bridge Priority/ID:   0x8000/ 00:D0:C0:F5:18:C0
  Port Priority/ID:     0x80/ 0x1D
  Message Age:          0/256 seconds (exactly 0  seconds)
  Maximum Age:          5120/256 seconds (exactly 20  seconds)
  Hello Time:           512/256 seconds (exactly 2  seconds)
  Forward Delay:        3840/256 seconds (exactly 15  seconds)
```

Root Bridge Selection Criteria

- At the beginning, all bridges assume they are the center of the universe and declare themselves as the Root Bridge, by placing its own BID in the Root BID field of the BPDU.

Bytes	Field
2	Protocol ID
1	Version
1	Message Type
1	Flags
8	Root ID
4	Cost of Path
8	Bridge ID
2	Port ID
2	Message Age
2	Maximum Age Time
2	Hello Time
2	Forward Delay

310P_127

When first booted, root ID = bridge ID.

Elect one Root Bridge
Lowest BID wins!

Bytes	Field
2	Protocol ID
1	Version
1	Message type
1	Flags
8	Root ID
4	Cost of path
8	Bridge ID
2	Port ID
2	Message age
2	Max age
2	Hello time
2	Forward delay

310P_126

Who is the root bridge?

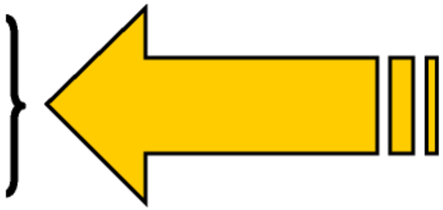
How far away is the root bridge?

What is the BID of the bridge that sent this BPDU?

What port on the sending bridge did this BPDU come from?

Local Switch Root Port Election

Bytes	Field
2	Protocol ID
1	Version
1	Message Type
1	Flags
8	Root ID
4	Cost of Path
8	Bridge ID
2	Port ID
2	Message Age
2	Maximum Age Time
2	Hello Time
2	Forward Delay



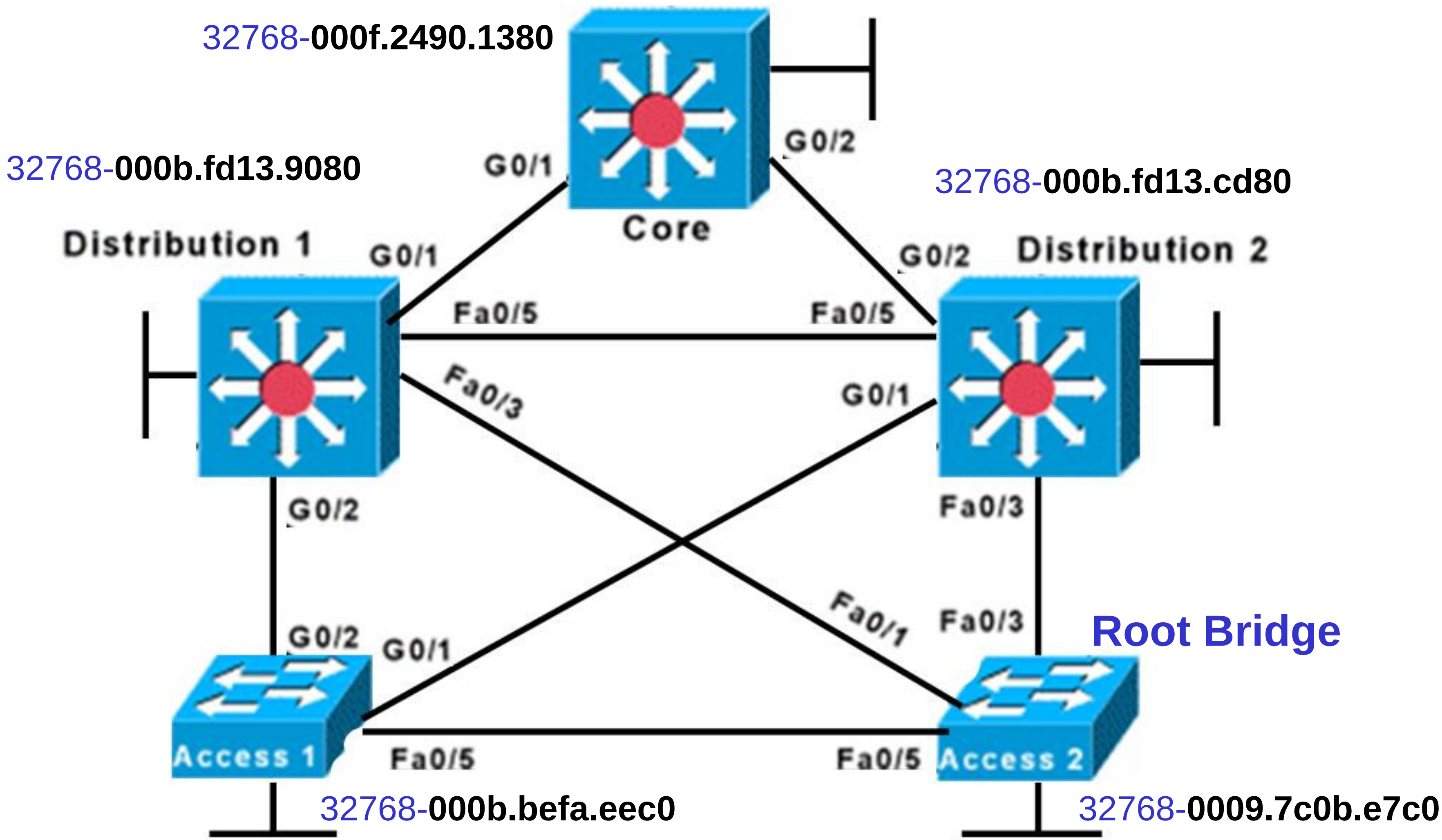
Lowest

- Path cost to root
- Sender BID
- Sender port ID

310P_002

What is the shortest path to the root bridge?

- Once all of the switches see that Access2 has the lowest BID, they are all in agreement that Access2 is the Root Bridge.



Elect Root Ports

STP Convergence

Step 1 Elect one Root Bridge

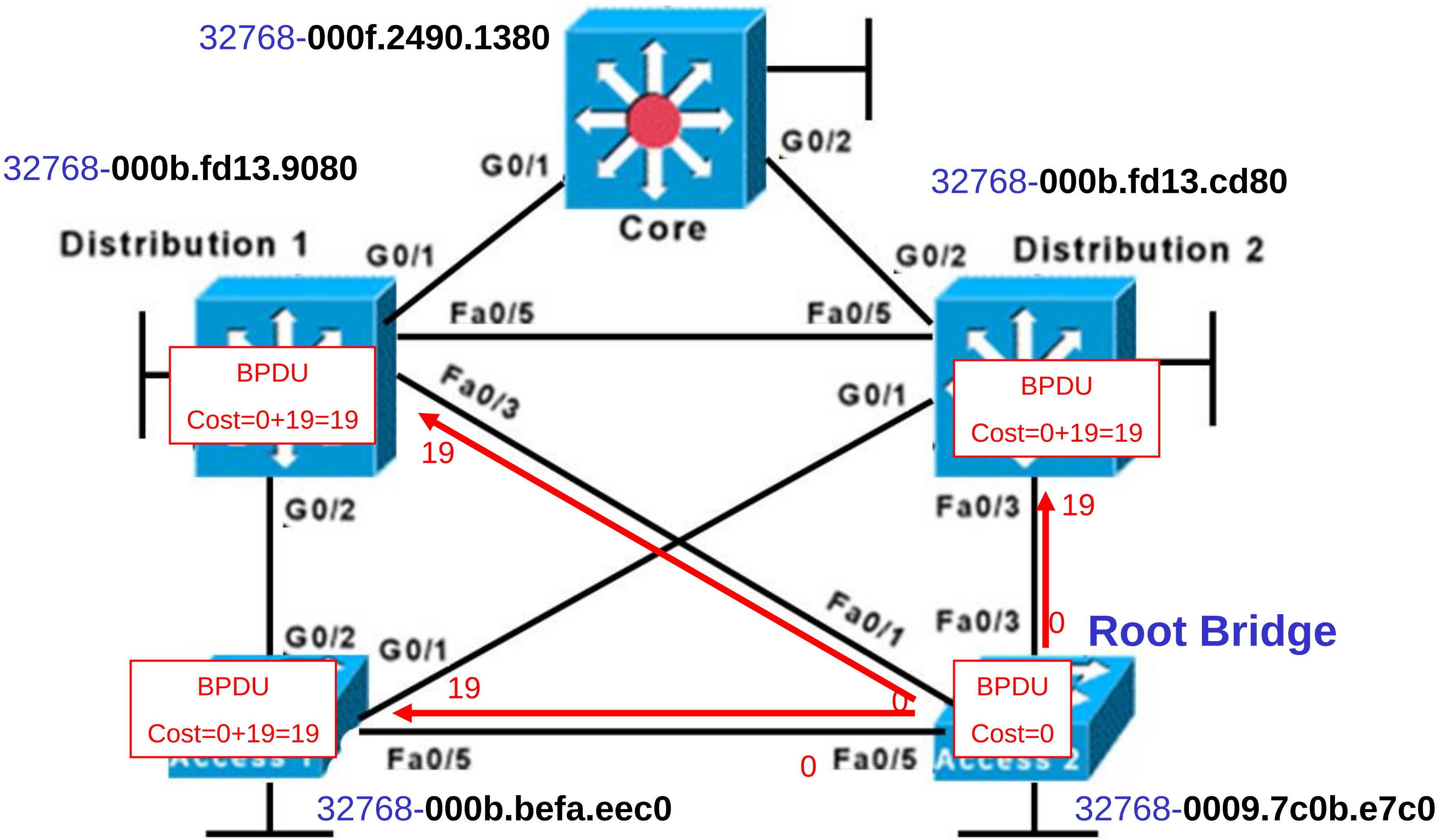
Step 2 Elect Root Ports

Step 3 Elect Designated Ports

Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100

- Now that the Root War has been won, switches move on to selecting **Root Ports**.
- A bridge's **Root Port** is the port closest to the Root Bridge.
- Bridges use the **cost** to determine closeness.
- **Every non-Root Bridge will select one Root Port!**
- Specifically, bridges track the **Root Path Cost**, the cumulative cost of all links to the Root Bridge.

- Root Bridge, Access2 sends out BPDUs, containing a Root Path Cost of 0.
- Access1, Distribution1, and Distribution2 receives these BPDUs and adds the Path Cost of the FastEthernet interface to the Root Path Cost contained in the BPDU.
- Access1, Distribution1, and Distribution2 add Root Path Cost 0 PLUS its Port cost of 19 = 19.
- This value is used internally and used in BPDUs to other switches..



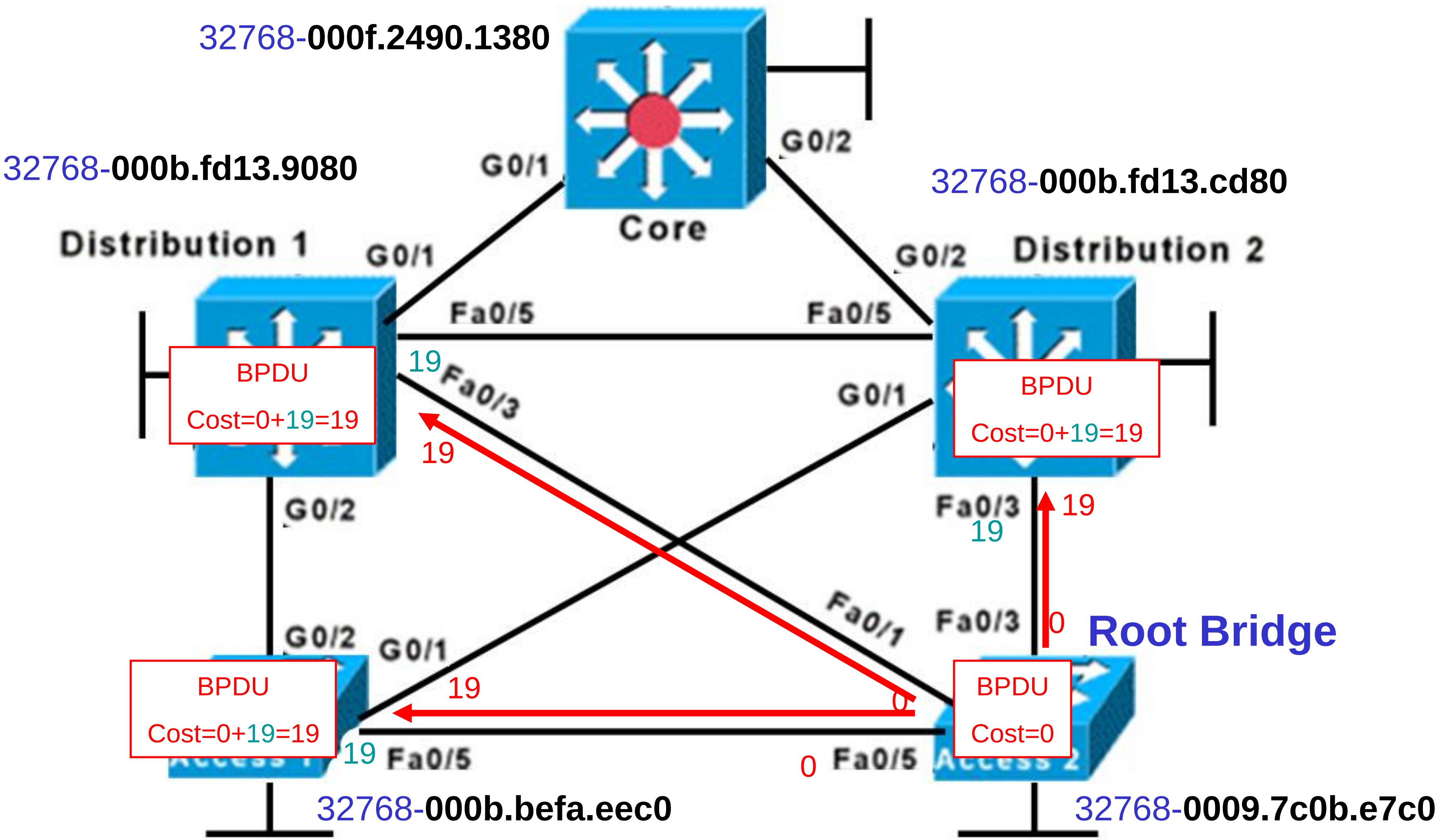
Difference b/t Path Cost and Root Path Cost

Path Cost:

- The value assigned to each port.
- Added to BPDUs received on that port to calculate Root Path Cost.

Root Path Cost

- Cumulative cost to the Root Bridge.
- This is the value transmitted in the BPDU.
- Calculated by adding the receiving port's Path Cost to the valued contained in the BPDU.



show spanning-tree

Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100

```
Distribution1#show spanning-tree
VLAN0001
  Spanning tree enabled protocol ieee
  Root ID     Priority     32769
              Address      0009.7c0b.e7c0
              Cost         19
              Port         3 (FastEthernet0/3)
              Hello Time    2 sec  Max Age 20 sec  Forward Delay 15 sec
  Bridge ID   Priority     32769  (priority 32768 sys-id-ext 1)
              Address      000b.fd13.9080
              Hello Time    2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time 300
Interface        Port ID                     Designated                Port ID
Name             Prio.Nbr      Cost Sts      Cost Bridge ID            Prio.Nbr
---------------- -------- --------- --- --------- -------------------- --------
Fa0/1            128.1           19 BLK        19 32769 000b.befa.eec0 128.1
Fa0/2            128.2           19 BLK        19 32769 000b.befa.eec0 128.2
Fa0/3            128.3           19 FWD         0 32769 0009.7c0b.e7c0 128.1
Fa0/4            128.4           19 BLK         0 32769 0009.7c0b.e7c0 128.2
Fa0/5            128.5           19 FWD        19 32769 000b.fd13.9080 128.5
Gi0/1            128.25           4 FWD        19 32769 000b.fd13.9080 128.25
Interface        Port ID                     Designated                Port ID
Name             Prio.Nbr      Cost Sts      Cost Bridge ID            Prio.Nbr
---------------- -------- --------- --- --------- -------------------- --------
Gi0/2            128.26           4 BLK        19 32769 000b.befa.eec0 128.26
```

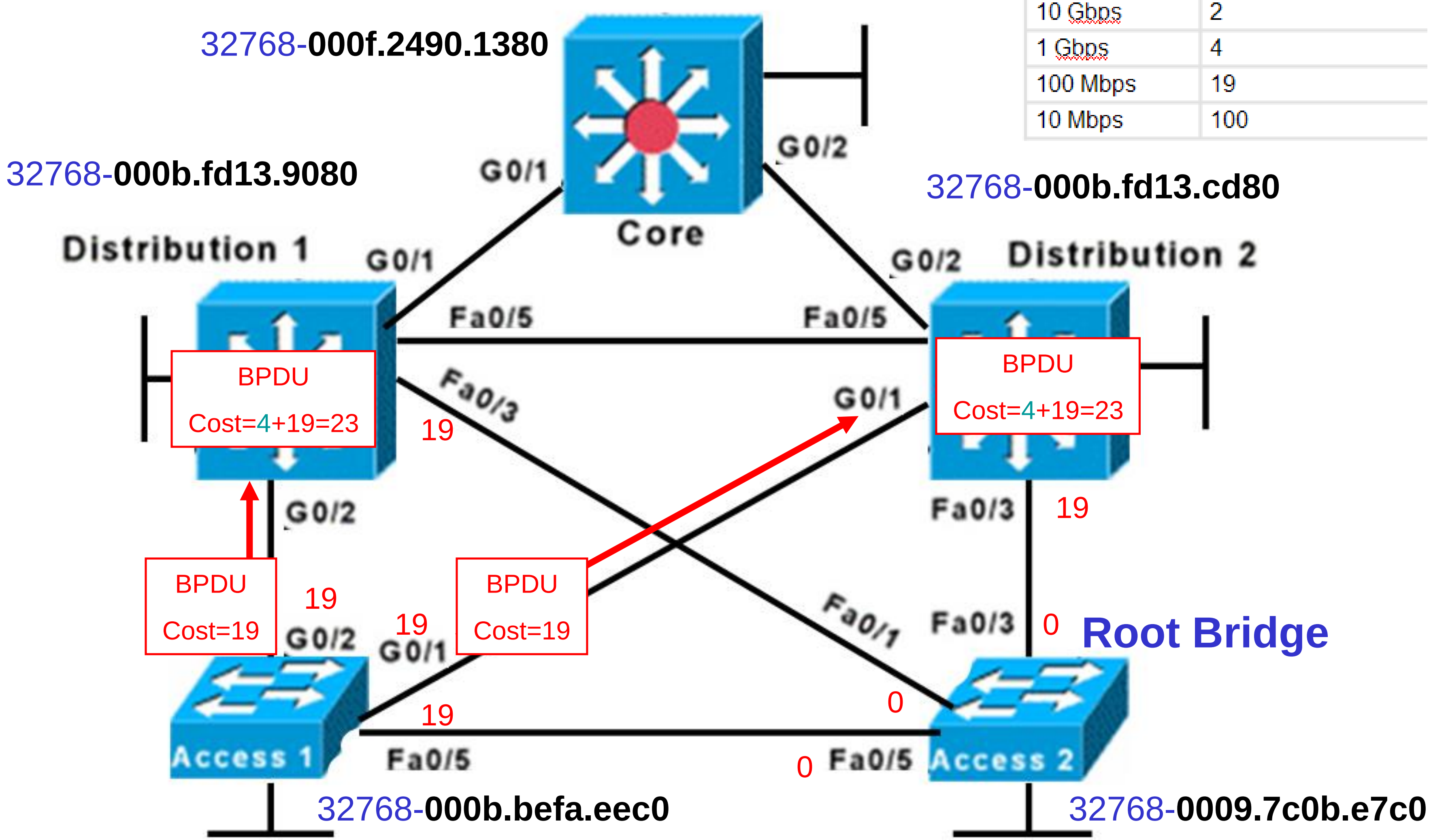
show spanning-tree detail

Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100

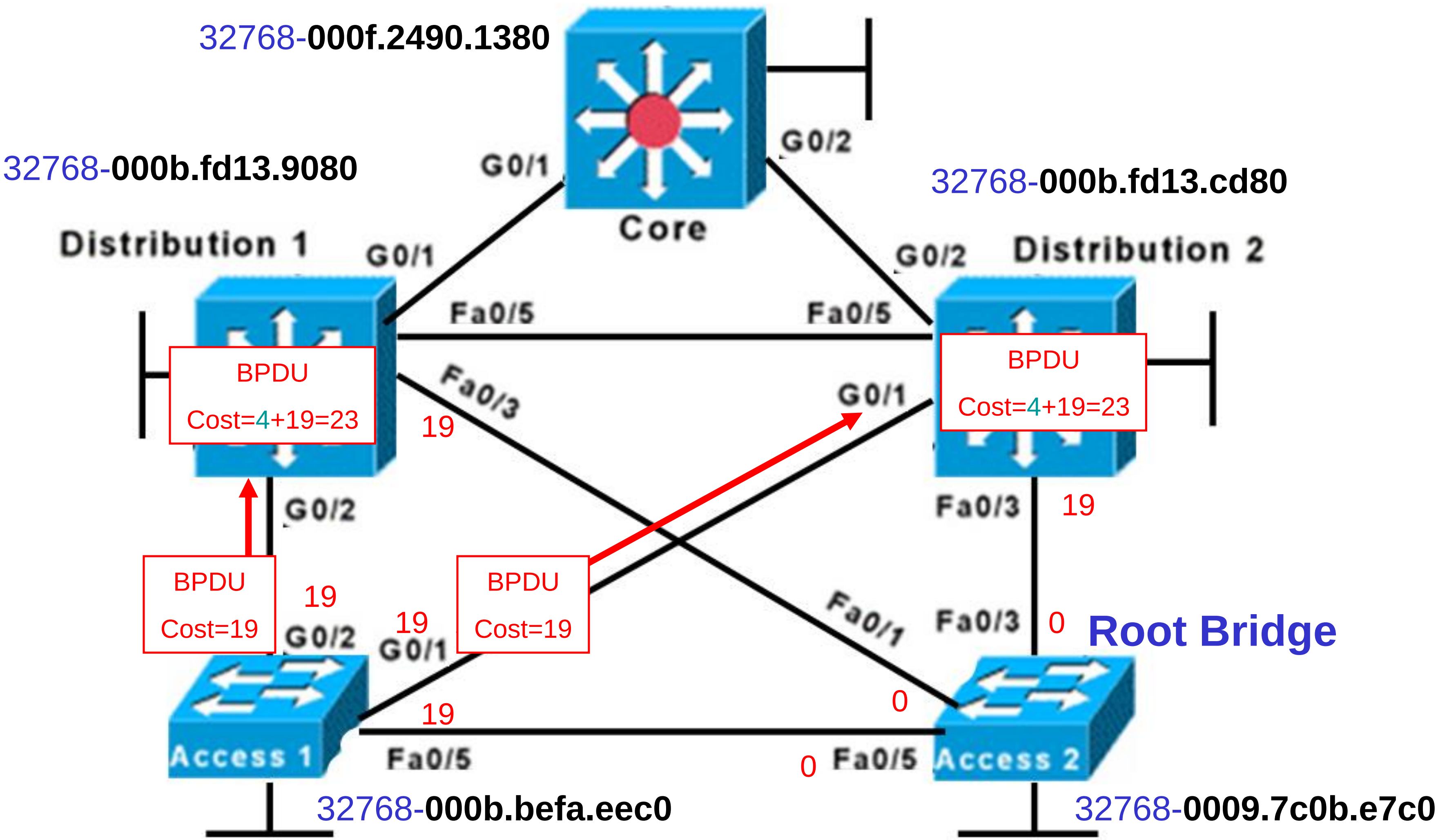
```
Distribution1#show spanning-tree detail
 VLAN0001 is executing the ieee compatible Spanning Tree protocol
  Bridge Identifier has priority 32768, sysid 1, address 000b.fd13.9080
  Configured hello time 2, max age 20, forward delay 15
  Current root has priority 32769, address 0009.7c0b.e7c0
  Root port is 3 (FastEthernet0/3), cost of root path is 19
  Topology change flag not set, detected flag not set
  Number of topology changes 7 last change occurred 00:14:34 ago
          from GigabitEthernet0/1
  Times:  hold 1, topology change 35, notification 2
          hello 2, max age 20, forward delay 15
  Timers: hello 0, topology change 0, notification 0, aging 300
```

- Switches now send BPDUs with their Root Path Cost out other interfaces.
- **Note**: STP costs are incremented as BPDUs are received on a port, not as they are sent out a port.
- Access 1 uses this value of 19 internally and sends BPDUs with a Root Path Cost of 19 out all other ports.

Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100

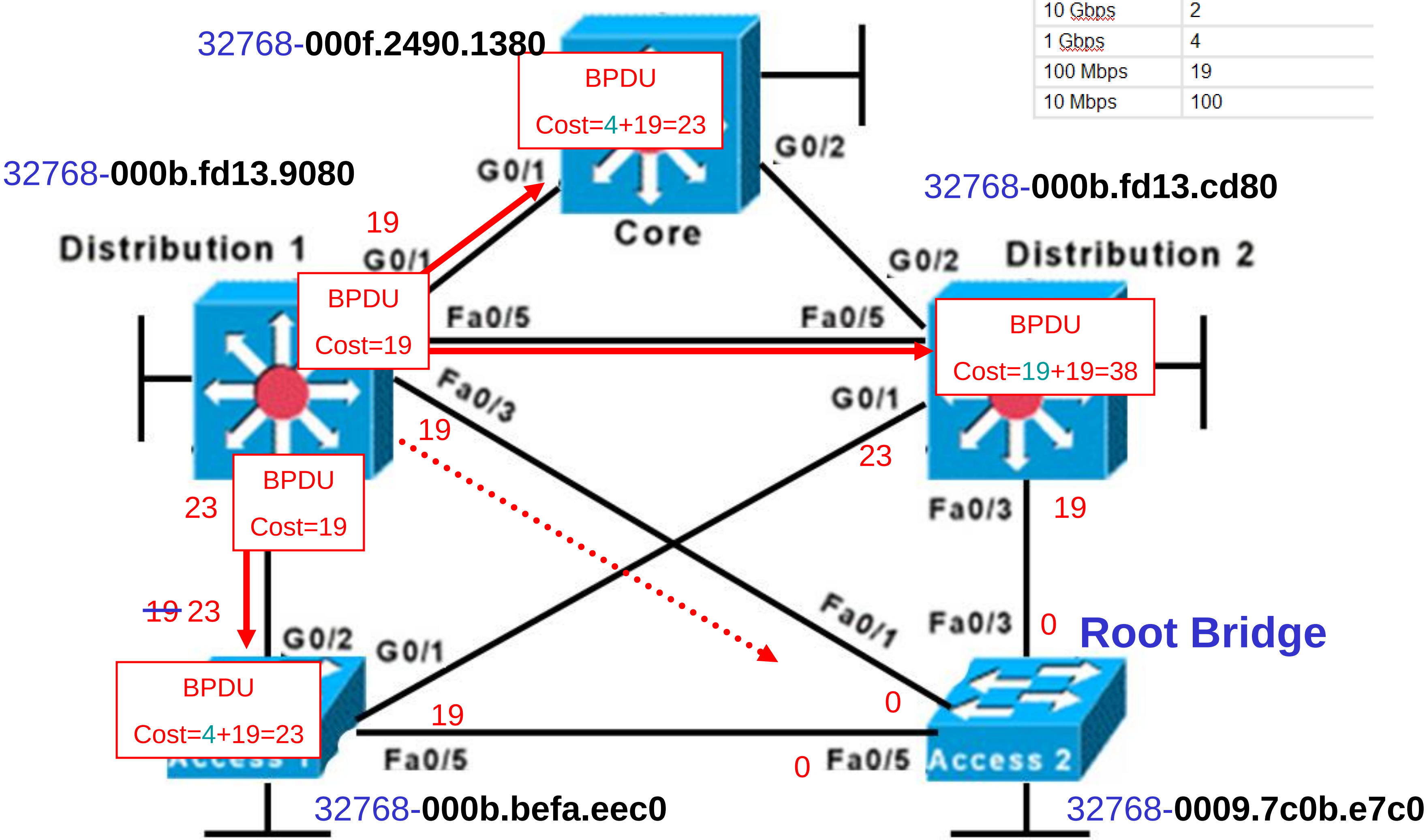


- Distribution 1 and Distribution 2 receive the BPDUs from Access 1, and adds the Path Cost of 4 to those interfaces, giving a Root Path Cost of 23.
- However, both of these switches already have an “internal” Root Path Cost of 19 that was received on another interface.
- Distribution 1 and Distribution 2 use the better BPDU of 19 when sending out their BPDUs to other switches.



- Distribution 1 now sends BPDUs with its Root Path Cost out other interfaces.
- Again, STP costs are incremented as BPDUs are received on a port, not as they are sent out a port.

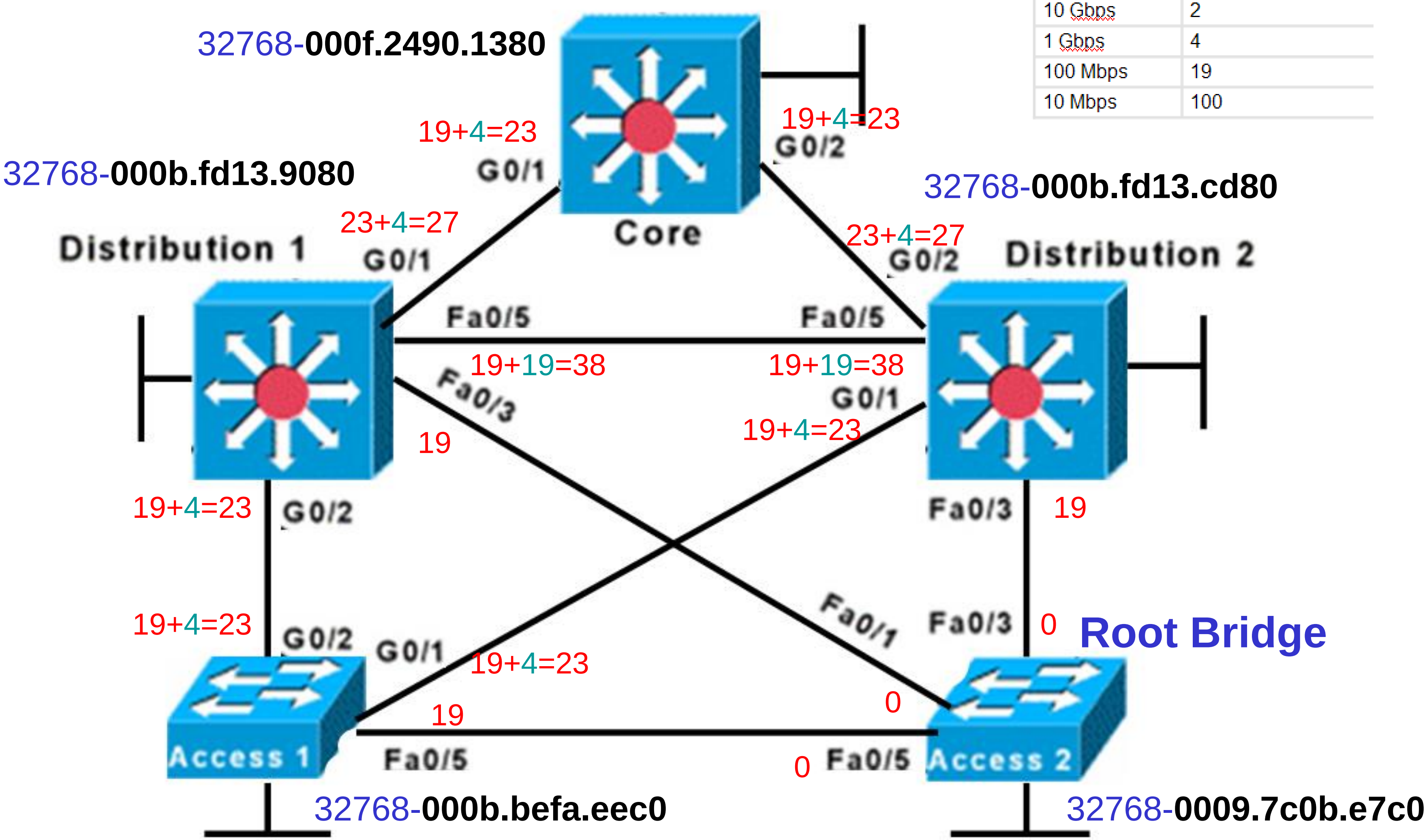
Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100



Final Results

- Ports show ***Received*** Root Path Cost = BPDU Root Path Cost + Path Cost of Interface, after the "best" BPDU is received on that port from the neighboring switch.
- This is the cost of reaching the Root Bridge from this interface towards the neighboring switch.
- ***Now let's see how this is used!***

Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100



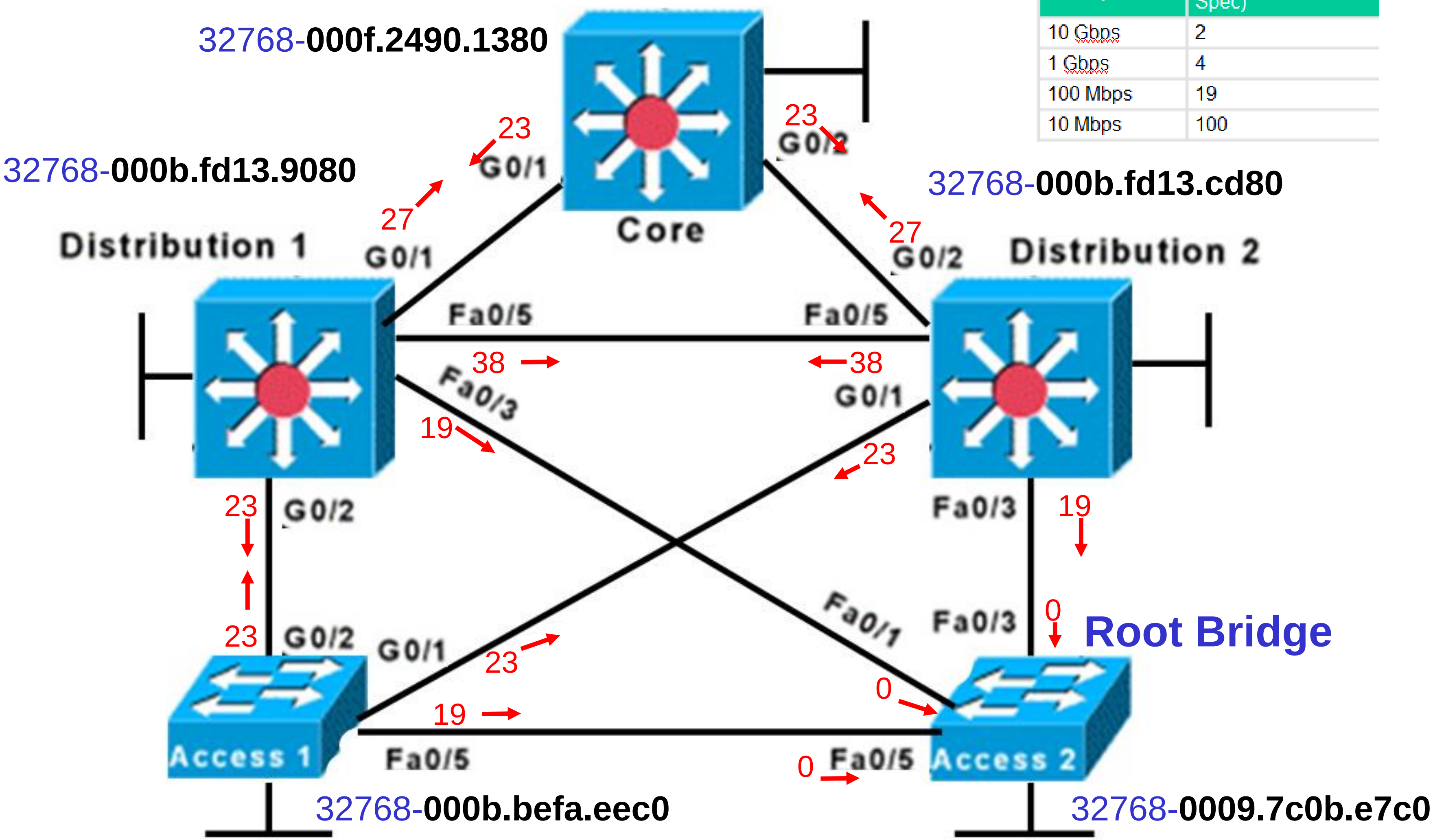
Next:

- Elect Root Ports
- Elect Designated Ports
- Non-Designated Ports: All other ports

Elect Root Ports

- Every non-Root bridge must select one **Root Port**.
- A bridge's **Root Port** is the port closest to the Root Bridge.
- Bridges use the **cost** to determine closeness.

Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100

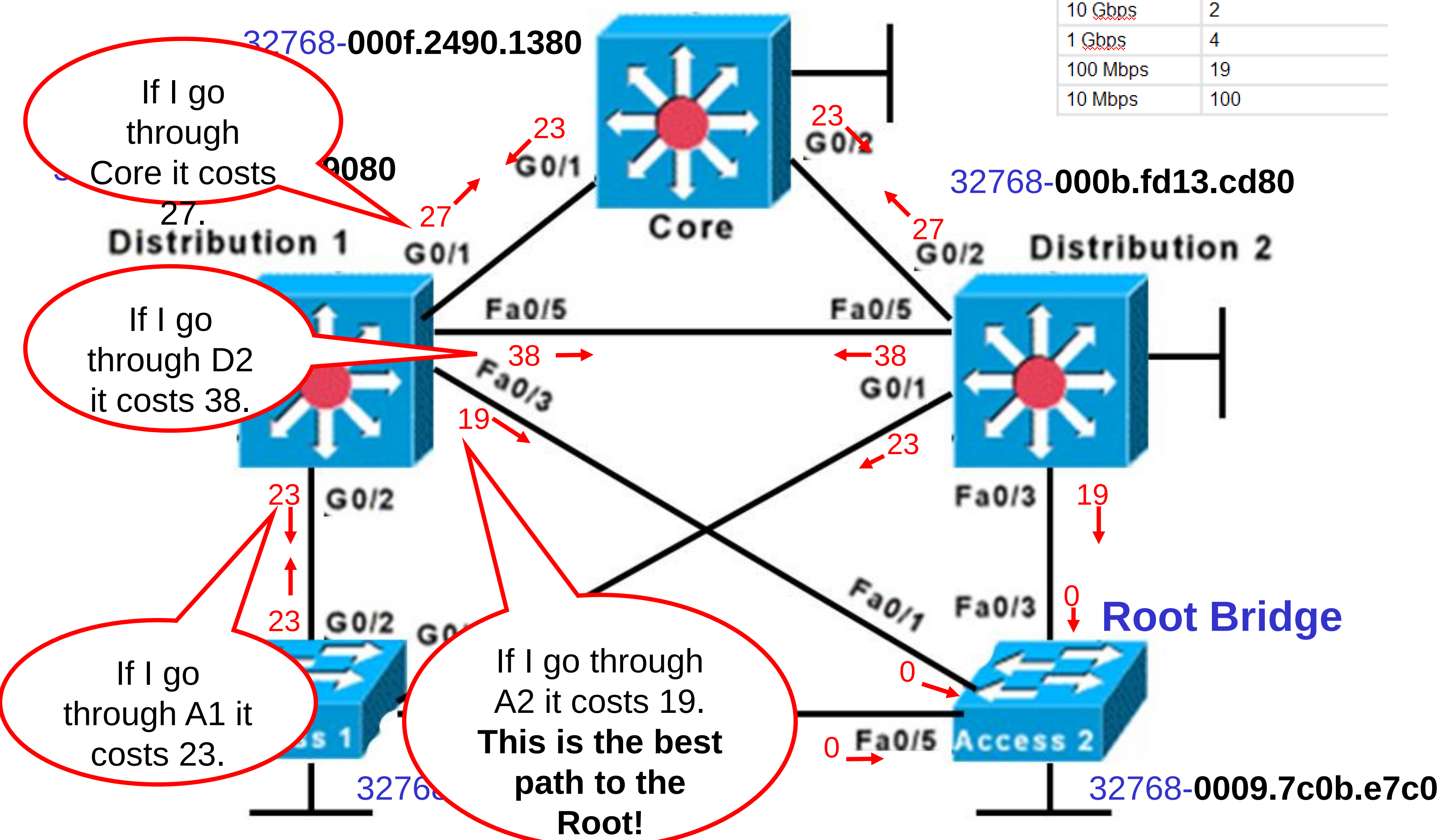


Elect Root Ports: (Review)

- Ports show ***Received*** Root Path Cost = BPDU Root Path Cost + Path Cost of Interface, after the “best” BPDU is received on that port from the neighboring switch.
- This is the cost of reaching the Root Bridge from this interface towards the neighboring switch.

Distribution 1 “thought process”

Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100

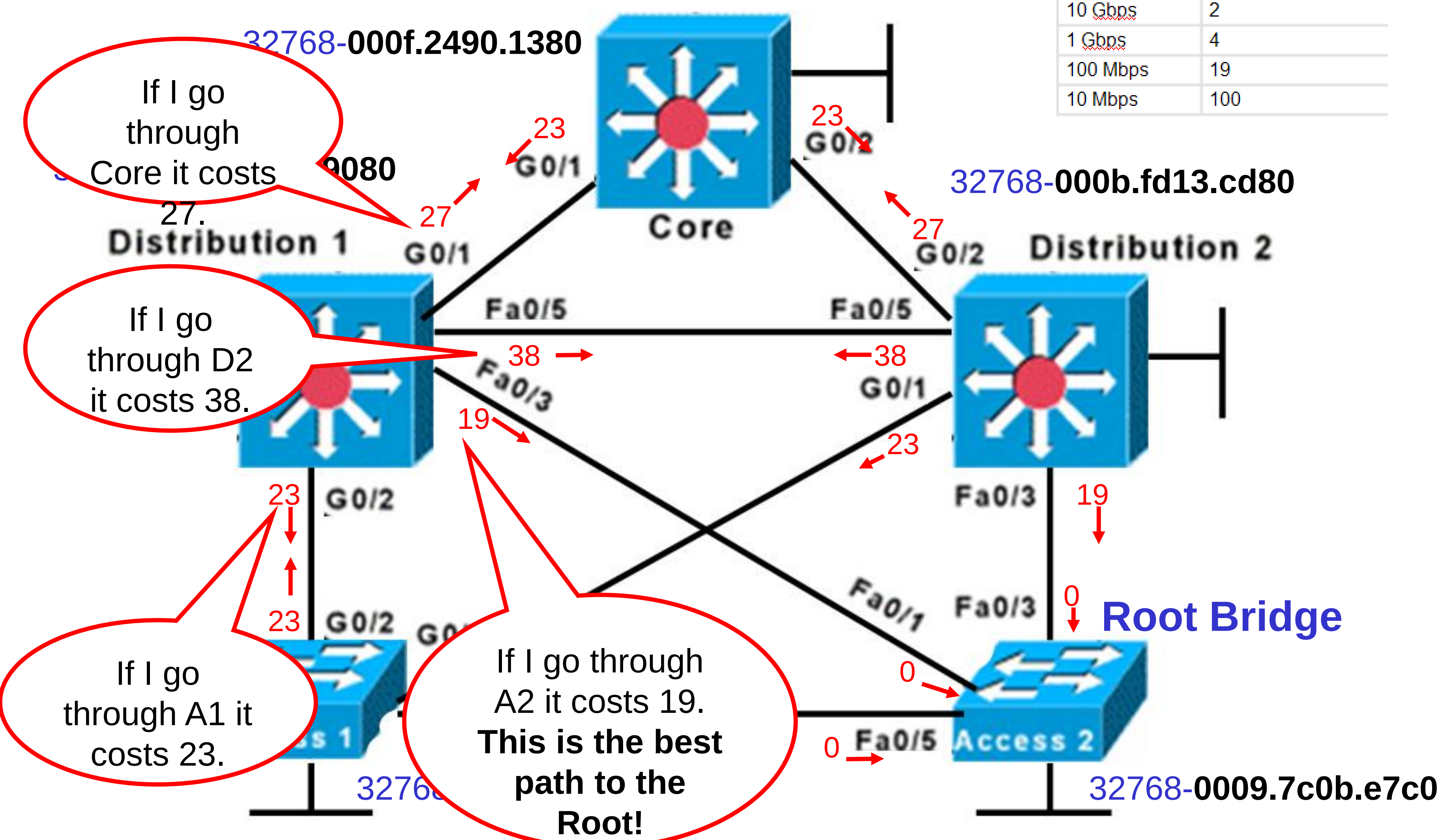


Elect Root Ports:

- This is from the ***switch's perspective***.
- Switch, "What is my cost to the Root Bridge?"
- Later we will look at Designated Ports, which is from the ***Segment's perspective***.

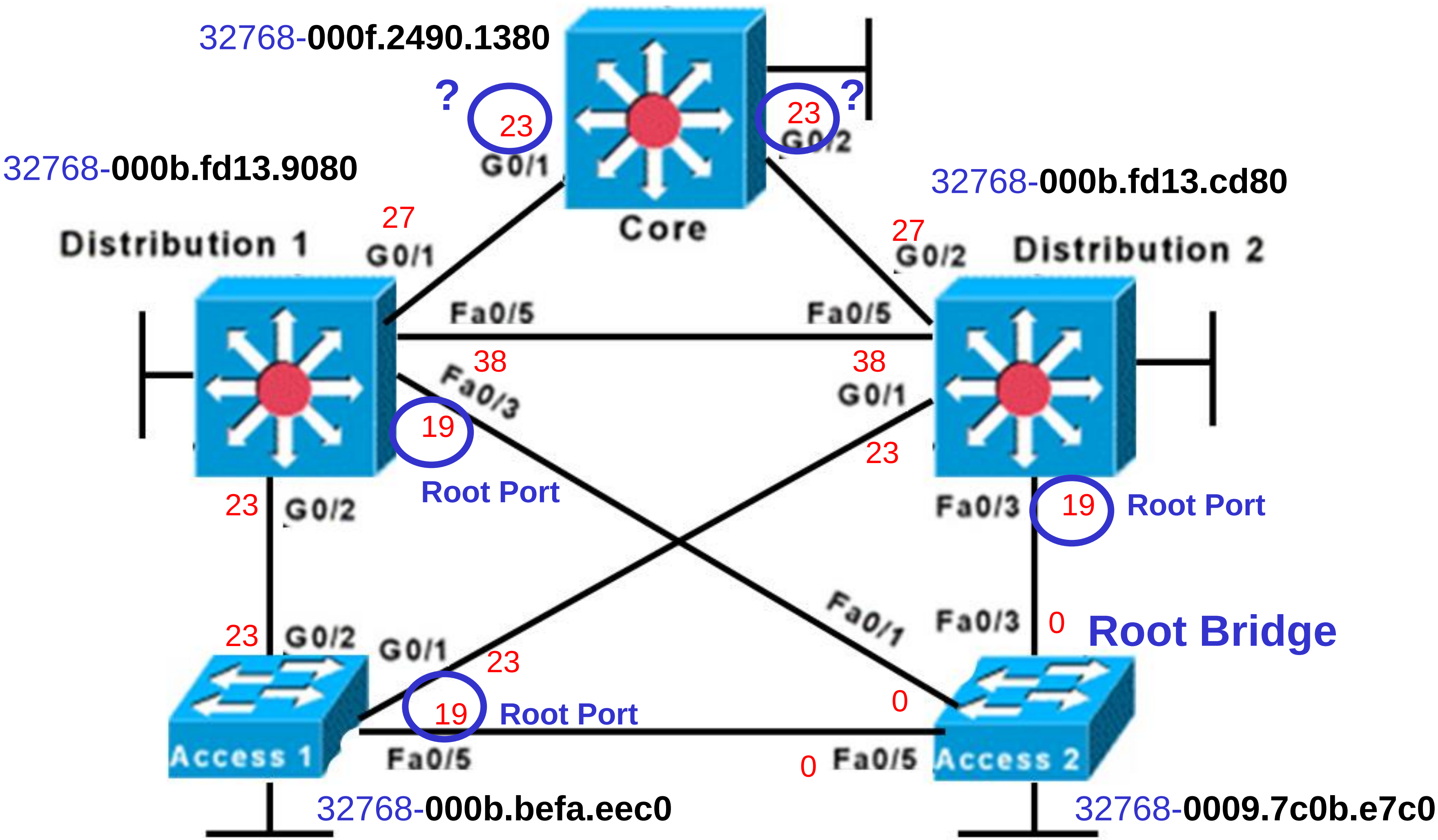
Distribution 1 "thought process"

Link Speed	Cost (Revised IEEE Spec)
10 Gbps	2
1 Gbps	4
100 Mbps	19
10 Mbps	100



Elect Root Ports

- Every non-Root bridge must select one **Root Port**.
- A bridge's **Root Port** is the port closest to the Root Bridge.
- Bridges use the **cost** to determine closeness.



Elect Root Ports

- Core switch has two equal Root Path Costs to the Root Bridge.
- In this case we need to look at the five-step decision process.

Five-Step decision Sequence

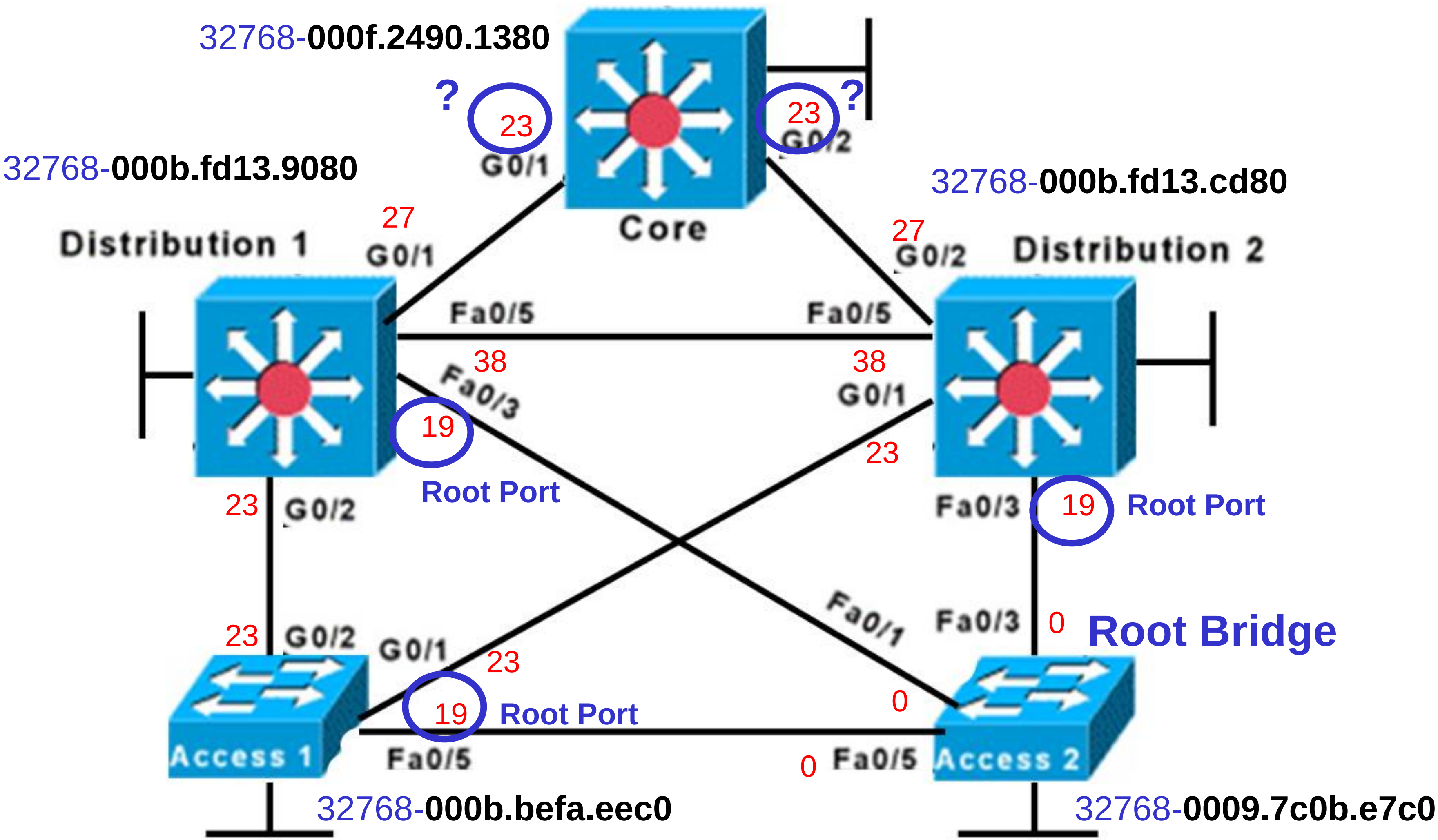
Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 - Lowest Port Priority

Step 5 - Lowest Port ID

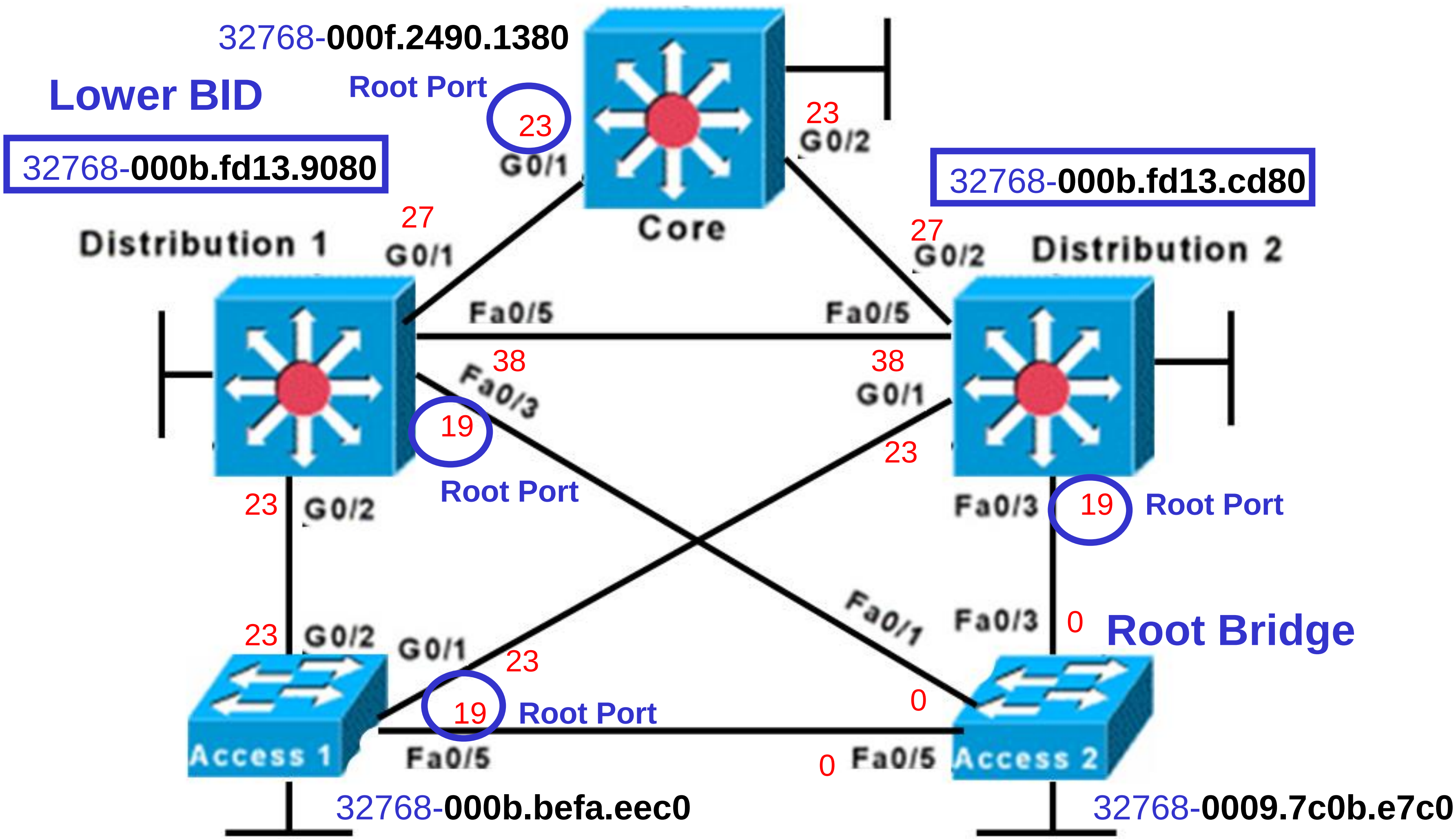


Elect Root Ports

- Distribution 1 switch has a lower Sender BID than Distribution 2.
- Core chooses the Root Port of G 0/1.

Five-Step decision Sequence

Step 1 - Lowest BID
Step 2 - Lowest Path Cost to Root Bridge
Step 3 - Lowest Sender BID
Step 4 - Lowest Port Priority
Step 5 - Lowest Port ID



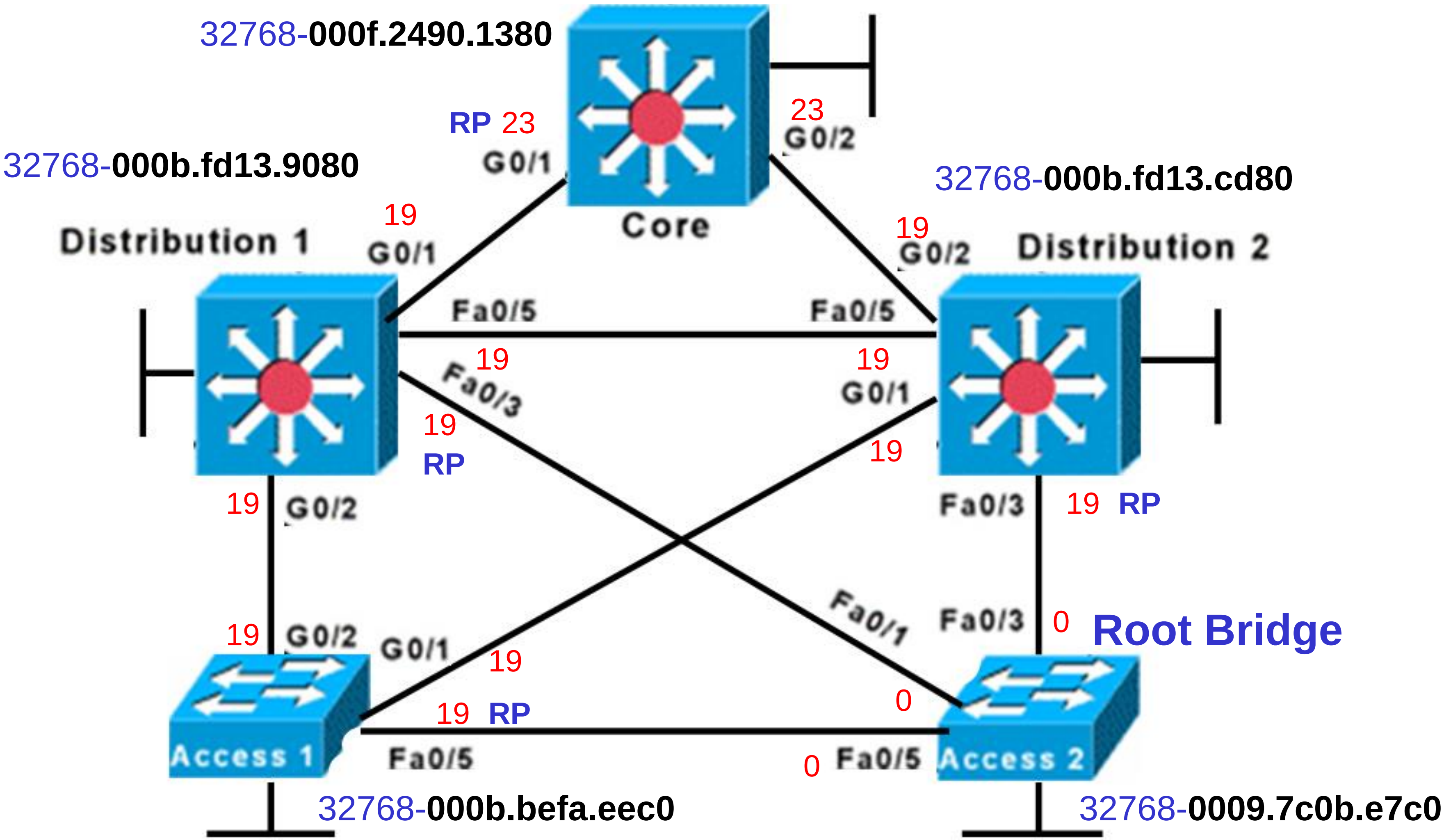
Elect Designated Ports

STP Convergence

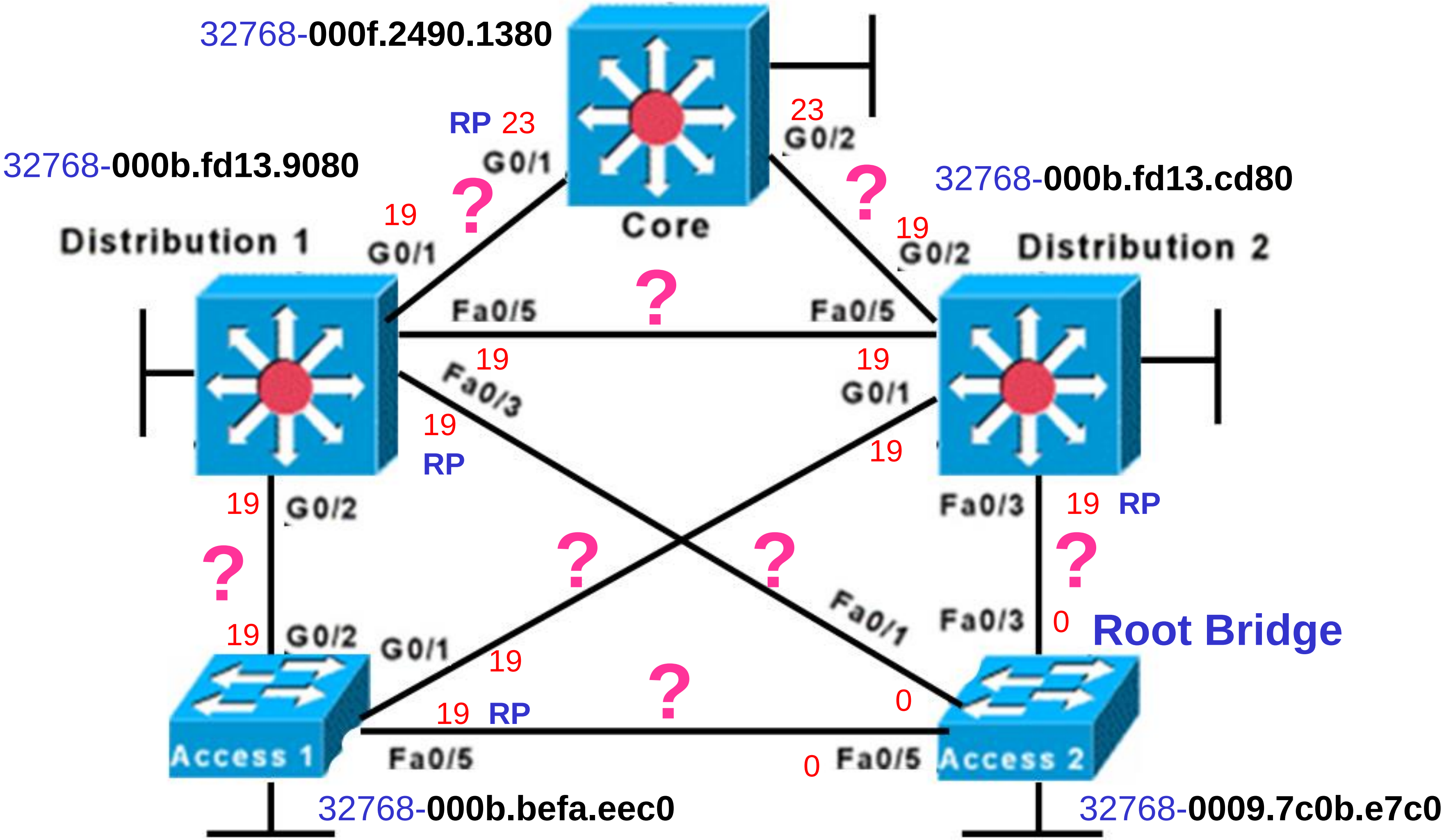
Step 1 Elect one Root Bridge
Step 2 Elect Root Ports
Step 3 Elect Designated Ports

- The loop prevention part of STP becomes evident during this step, electing designated ports.
- A Designated Port functions as the single bridge port that both sends and receives traffic to and from that segment and the Root Bridge.
- **Each segment in a bridged network has one Designated Port, chosen based on cumulative Root Path Cost to the Root Bridge.**
- The switch containing the **Designated Port** is referred to as the **Designated Bridge** for that segment.
- To locate Designated Ports, lets take a look at each segment.
- **Segment's perspective**: From a device on this segment, *"Which switch should I go through to reach the Root Bridge?"*
  - **Root Path Cost**, the cumulative cost of all links to the Root Bridge.
  - Obviously, the segment has not ability to make this decision, so the perspective and the decision is that of the switches on that segment.

- A **Designated Port** is elected for every segment.
- The **Designated Port** is the only port that sends and receives traffic to/from that segment to the Root Bridge, the best port towards the root bridge.
- **Note**: The Root Path Cost shows the ***Sent*** Root Path Cost.
- This is the advertised cost in the BPDU, by this switch out that interface, i.e. this is the cost of reaching the Root Bridge through me!

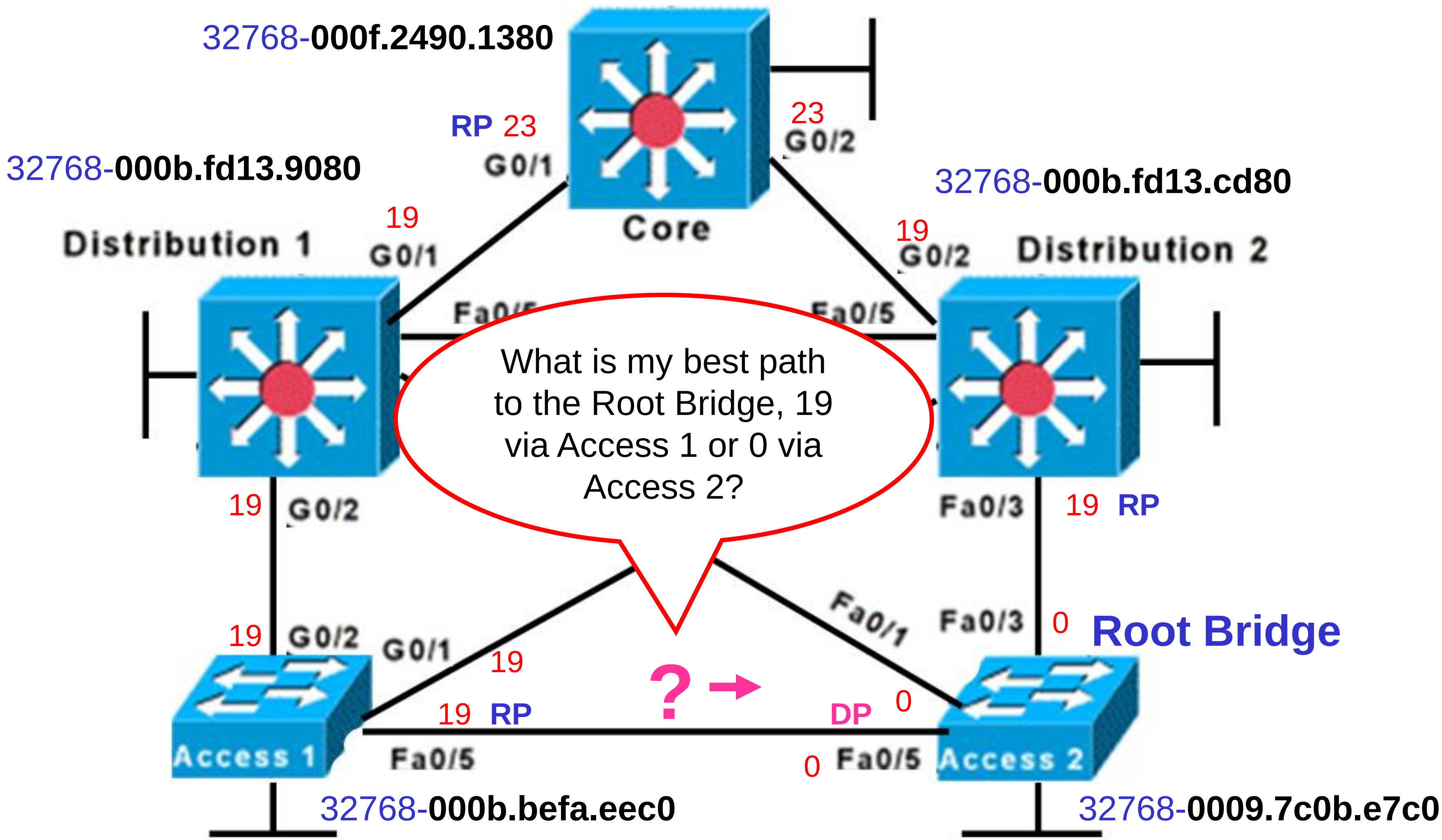


- A **Designated Port** is elected for every segment.
- **Segment's perspective**: From a device on this segment, *"Which switch should I go through to reach the Root Bridge?"*
- *"I'll decide using the advertised Root Path Cost from each switch!"*



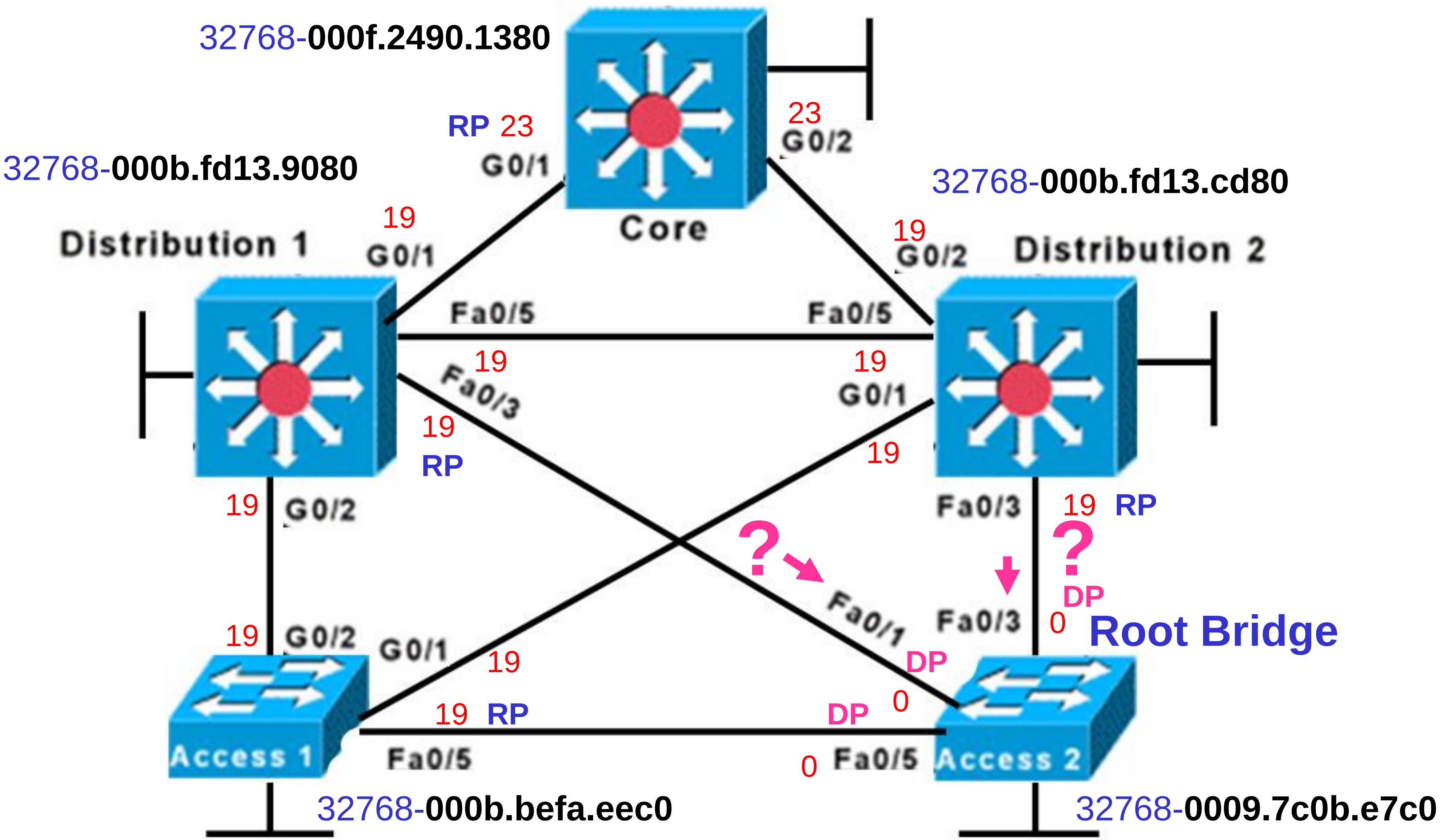
Segment's perspective:

- Access 2 has a Root Path Cost = 0 (after all it is the Root Bridge) and Access 1 has a Root Path Cost = 19.
- Because Access 2 has the lower Root Path Cost it becomes the **Designated Port** for that segment.



Segment's perspective:

- The same occurs between Access 2 and Distribution 1 and Distribution 2 switches.
- Because Access 2 has the lower Root Path Cost it becomes the **Designated Port** for those segments.

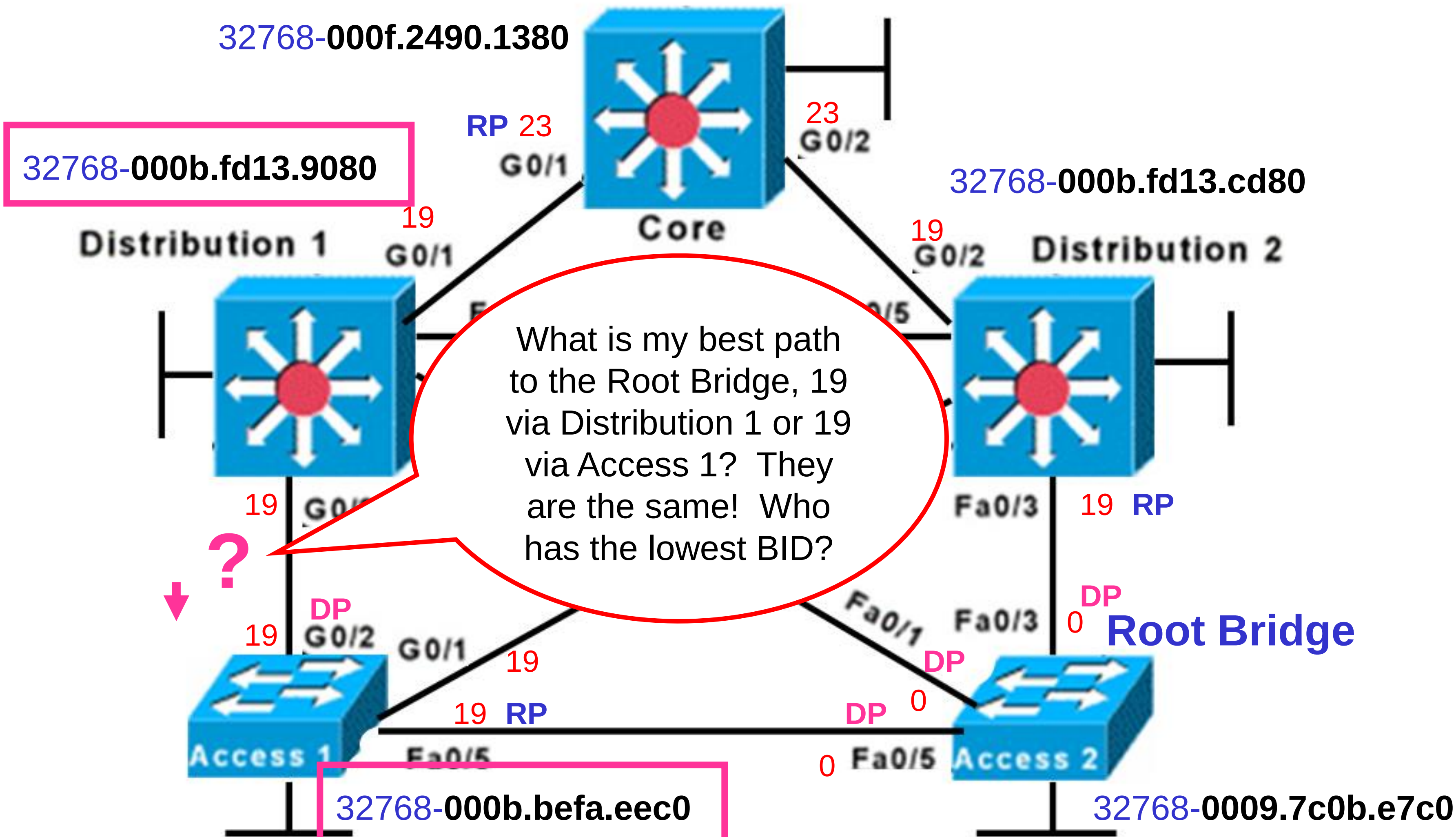


Segment's perspective:

- Segment between Distribution 1 and Access 1 has two equal Root Path Costs of 19.
- Using the Lowest Sender ID (first two steps are equal), **Access 1** becomes the best path and the **Designated Port**.

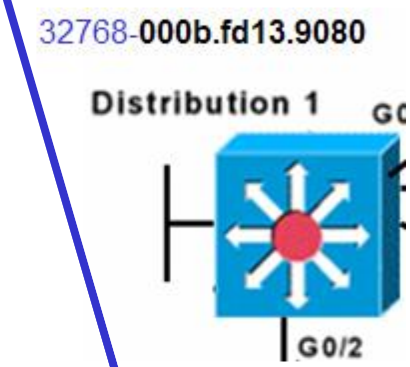
Five-Step decision Sequence

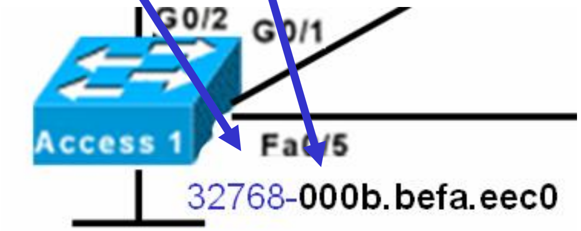
Step 1 - Lowest BID
Step 2 - Lowest Path Cost to Root Bridge
Step 3 - Lowest Sender BID
Step 4 - Lowest Port Priority
Step 5 - Lowest Port ID



Access 1 has Lower Sender BID


```
Distribution1#show spanning-tree detail
 Port 26 (GigabitEthernet0/2) of VLAN0001 is blocking
   Port path cost 4, Port priority 128, Port Identifier 128.26.
   Designated root has priority 32769, address 0009.7c0b.e7c0
   Designated bridge has priority 32769, address 000b.befa.eec0
   Designated port id is 128.26, designated path cost 19
   Timers: message age 3, forward delay 0, hold 0
   Number of transitions to forwarding state: 0
   BPDU: sent 2, received 1070

Access1#show spanning-tree detail
 Port 26 (GigabitEthernet0/2) of VLAN0001 is forwarding
   Port path cost 4, Port priority 128, Port Identifier 128.26.
   Designated root has priority 32769, address 0009.7c0b.e7c0
   Designated bridge has priority 32769, address 000b.befa.eec0
   Designated port id is 128.26, designated path cost 19
   Timers: message age 0, forward delay 0, hold 0
   Number of transitions to forwarding state: 1
   BPDU: sent 2243, received 1
```





Segment's perspective:

- Segment between Distrib. 1 and Distrib. 2 has two equal Root Path Costs of 19.
- Using the Lowest Sender ID (first two steps are equal), **Distribution 1** becomes the best path and the **Designated Port**.

Five-Step decision Sequence

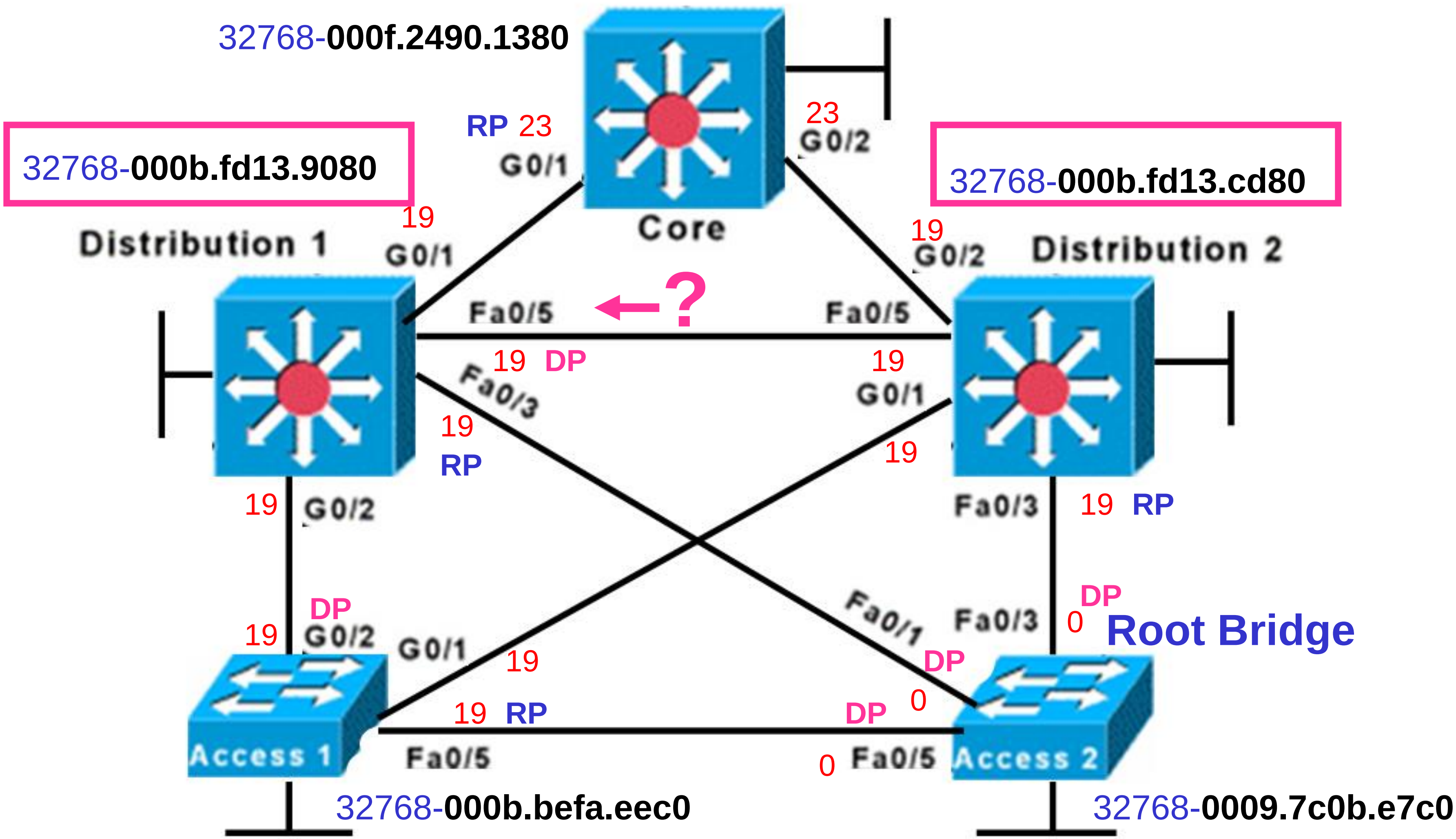
Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

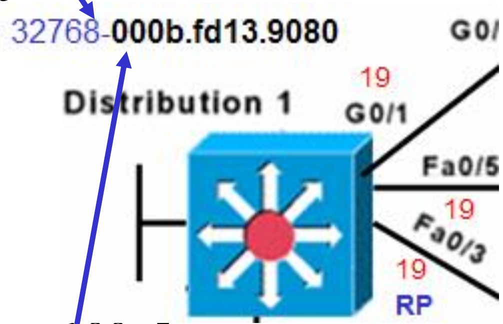
Step 4 - Lowest Port Priority

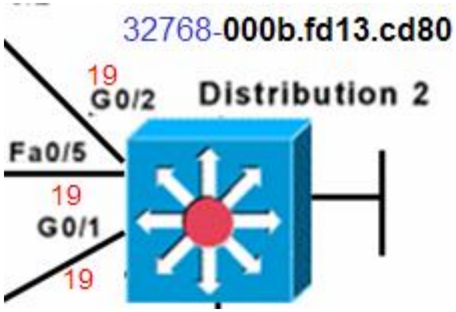
Step 5 - Lowest Port ID



Distribution 1 has Lower Sender BID


```
Distribution1#show spanning-tree detail
 Port 5 (FastEthernet0/5) of VLAN0001 is forwarding
   Port path cost 19, Port priority 128, Port Identifier 128.5.
   Designated root has priority 32769, address 0009.7c0b.e7c0
   Designated bridge has priority 32769, address 000b.fd13.9080
   Designated port id is 128.5, designated path cost 19
   Timers: message age 0, forward delay 0, hold 0
   Number of transitions to forwarding state: 1
   BPDU: sent 1074, received 0

Distribution2#show spanning-tree detail
 Port 5 (FastEthernet0/5) of VLAN0001 is blocking
   Port path cost 19, Port priority 128, Port Identifier 128.5.
   Designated root has priority 32769, address 0009.7c0b.e7c0
   Designated bridge has priority 32769, address 000b.fd13.9080
   Designated port id is 128.5, designated path cost 19
   Timers: message age 2, forward delay 0, hold 0
   Number of transitions to forwarding state: 0
   BPDU: sent 0, received 1097
```





Segment's perspective:

- Segment between Access 1 and Distrib. 2 has two equal Root Path Costs of 19.
- Using the Lowest Sender ID (first two steps are equal), **Access 1** becomes the best path and the **Designated Port**.

Five-Step decision Sequence

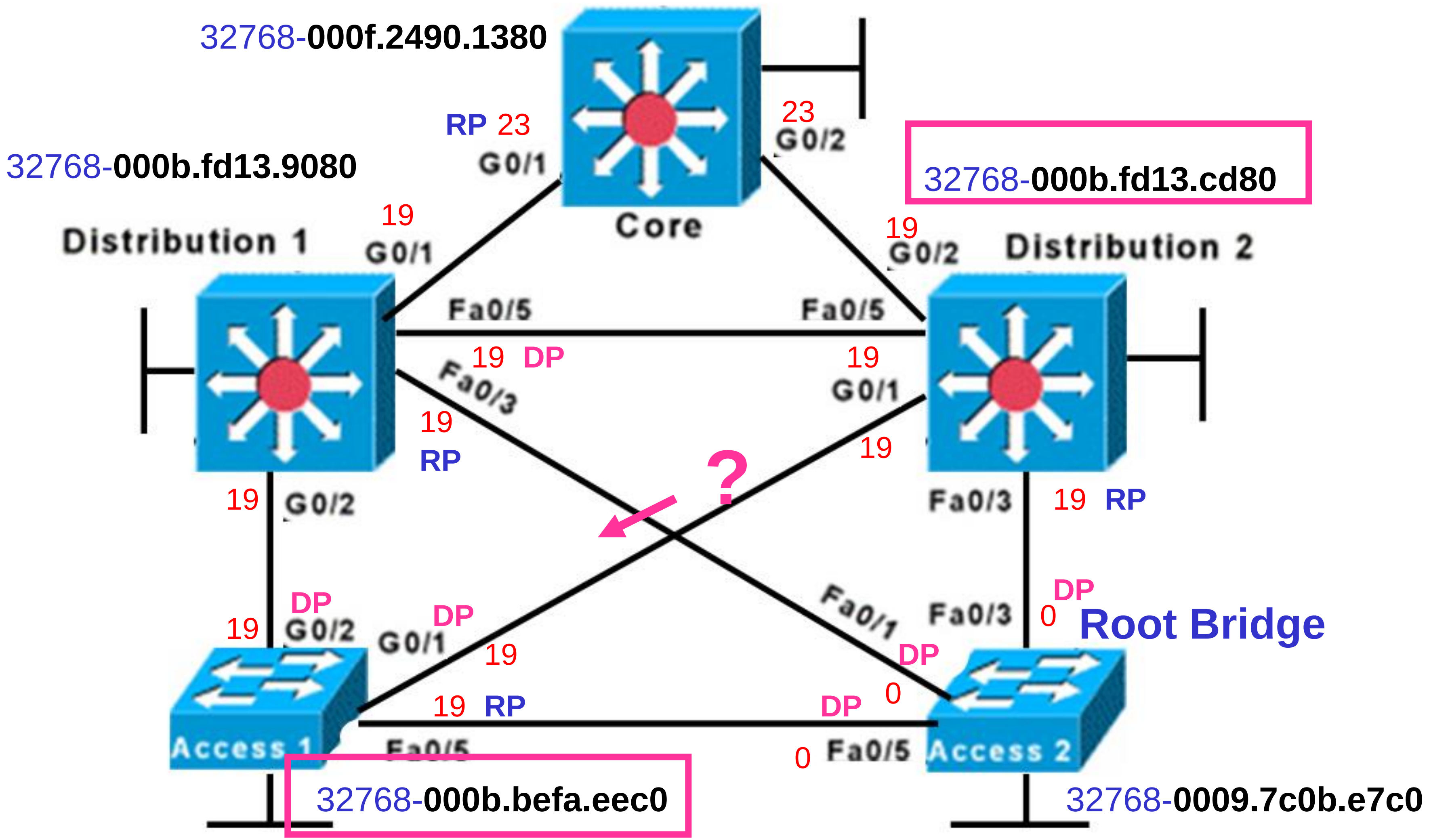
Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

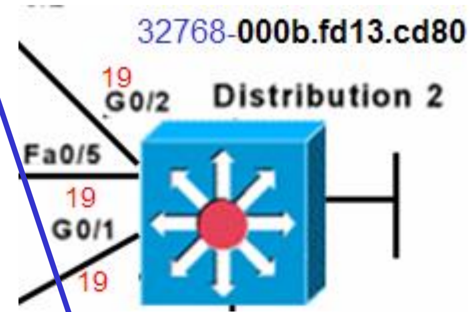
Step 4 - Lowest Port Priority

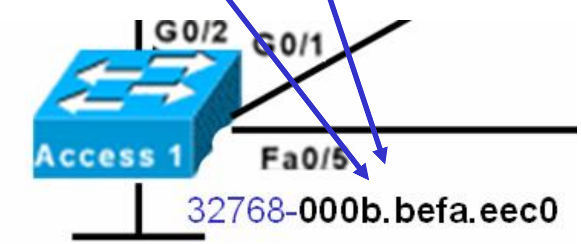
Step 5 - Lowest Port ID



Access 1 has Lower Sender BID

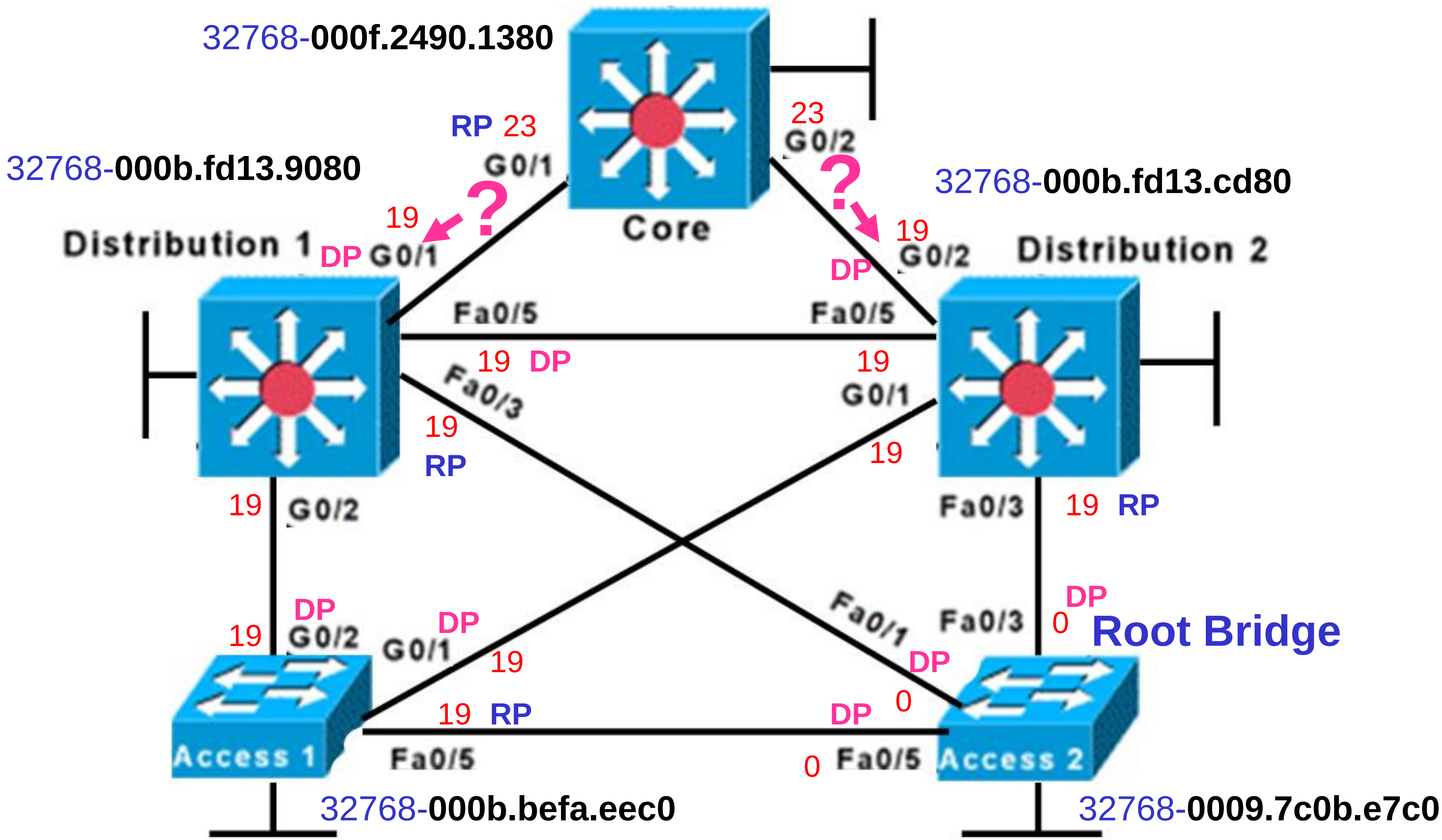

```
Distribution2#show spanning-tree detail
 Port 25 (GigabitEthernet0/1) of VLAN0001 is blocking
   Port path cost 4, Port priority 128, Port Identifier 128.25.
   Designated root has priority 32769, address 0009.7c0b.e7c0
   Designated bridge has priority 32769, address 000b.befa.eec0
   Designated port id is 128.25, designated path cost 19
   Timers: message age 3, forward delay 0, hold 0
   Number of transitions to forwarding state: 0
   BPDU: sent 2, received 1091

Access1#show spanning-tree detail
 Port 25 (GigabitEthernet0/1) of VLAN0001 is forwarding
   Port path cost 4, Port priority 128, Port Identifier 128.25.
   Designated root has priority 32769, address 0009.7c0b.e7c0
   Designated bridge has priority 32769, address 000b.befa.eec0
   Designated port id is 128.25, designated path cost 19
   Timers: message age 0, forward delay 0, hold 0
   Number of transitions to forwarding state: 1
   BPDU: sent 2240, received 1
```





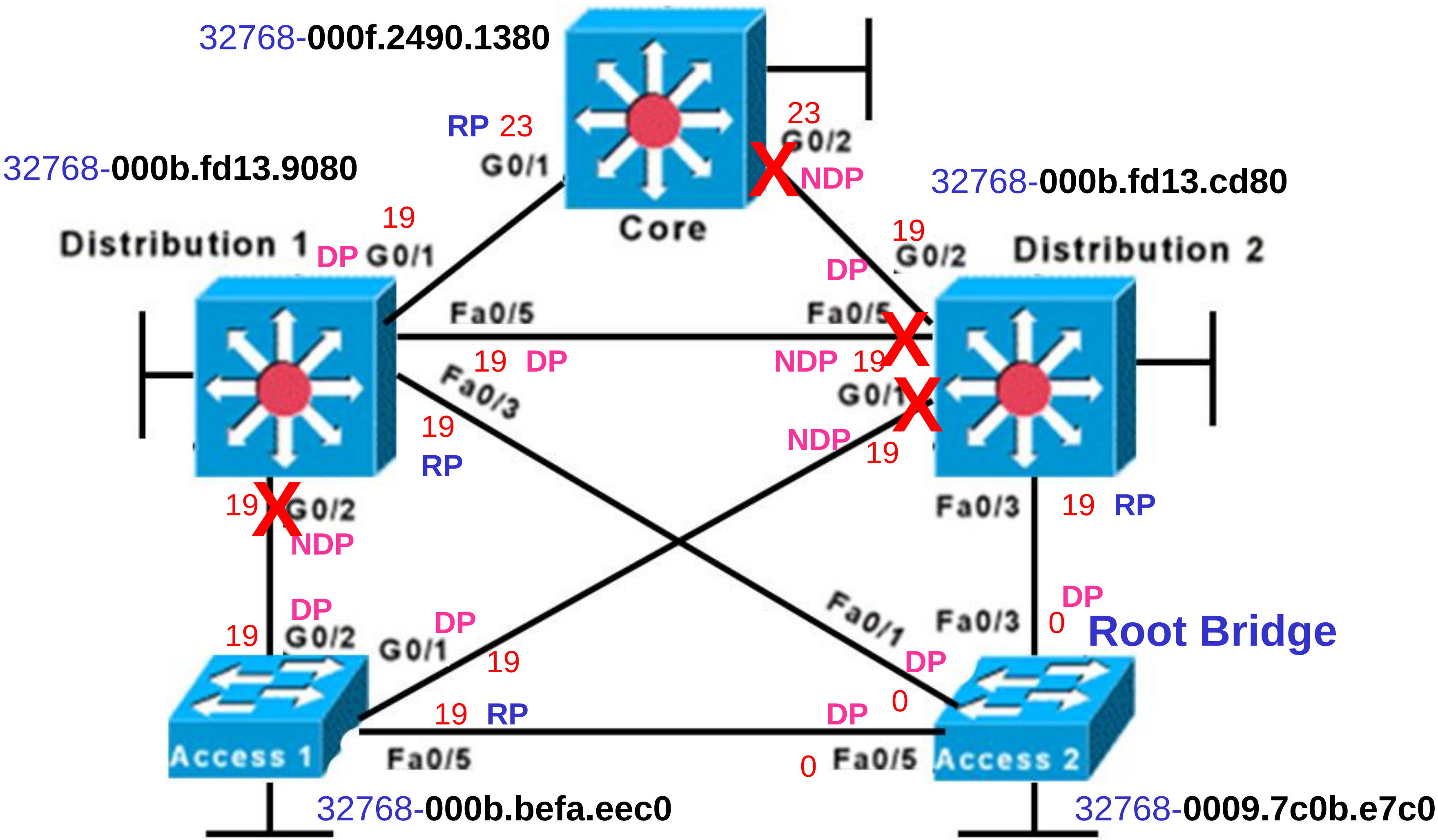
Segment's perspective:

- Because **Distribution 1** has the lower Root Path Cost it becomes the **Designated Port** for that segment.
- Because **Distribution 2** has the lower Root Path Cost it becomes the **Designated Port** for that segment.

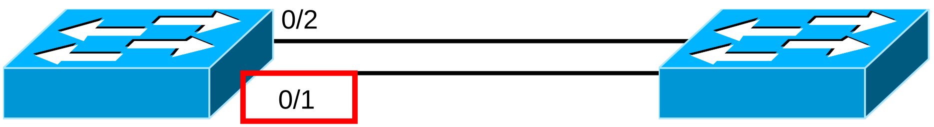


Segment's perspective:

- All other ports, those ports that are not Root Ports or Designated Ports, become **Non-Designated Ports.**
- **Non-Designated Ports** are put in blocking mode. (Coming)
- This is the loop prevention part of STP.



Port Cost/Port ID

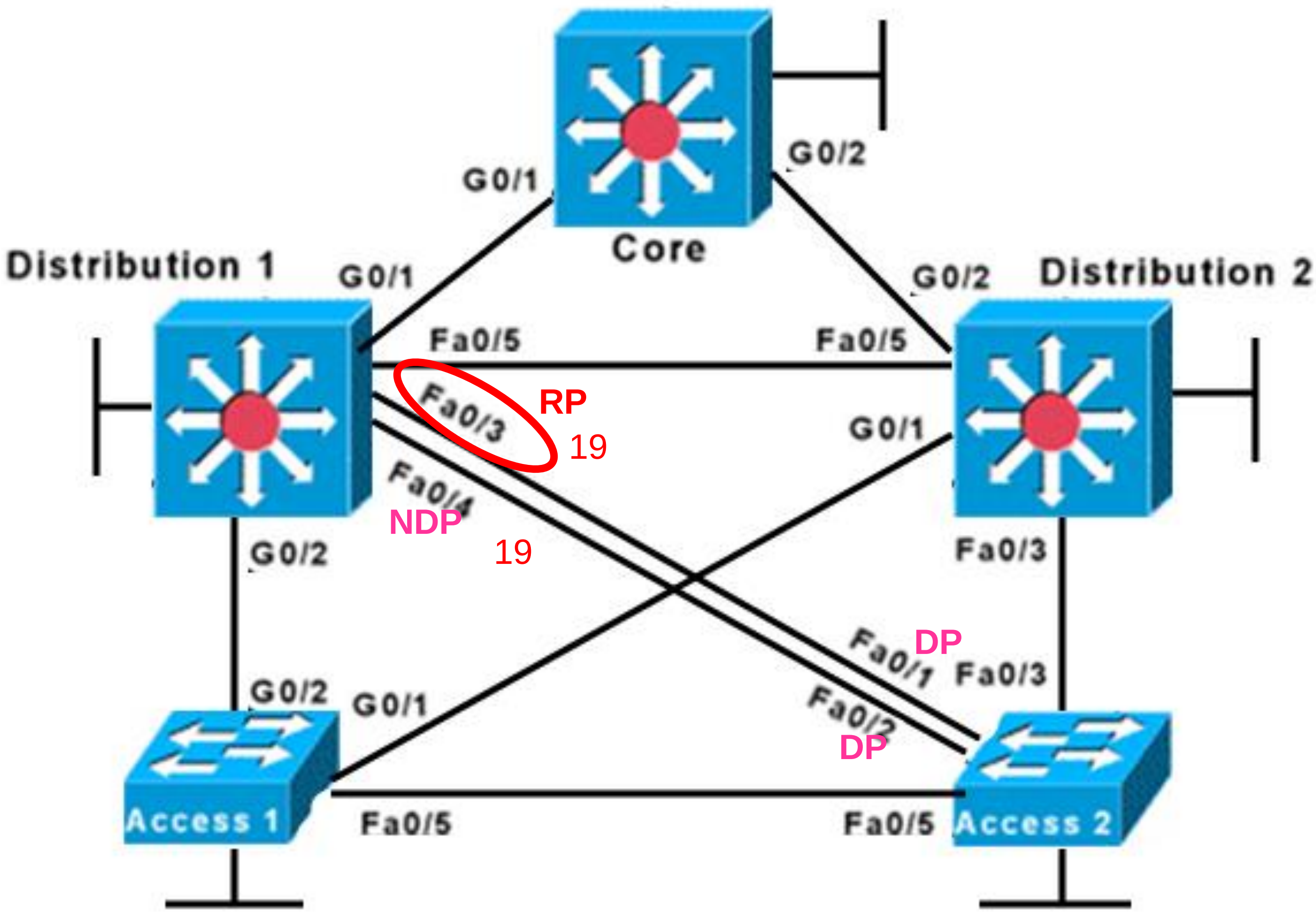


Assume path cost and port priorities are default (32). Port ID used in this case. Port 0/1 would forward because it's the lowest.

- If the path cost and bridge IDs are equal (as in the case of parallel links), the switch goes to the port priority as a tiebreaker.
- Lowest port priority wins (all ports set to 32).
- You can set the priority from 0 – 63.
- If all ports have the same priority, the port with the lowest port number forwards frames.

Port Cost/Port ID

- Fa 0/3 has a lower Port ID than Fa 04.
- More later (Fast EtherChannel)



Port Cost/Port ID

Cabrillo College

```
Distribution1#show spanning-tree
VLAN0001
  Spanning tree enabled protocol ieee
  Root ID    Priority    32769
             Address     0009.7c0b.e7c0
             Cost        19
             Port        3 (FastEthernet0/3)
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
  Bridge ID  Priority    32769  (priority 32768 sys-id-ext 1)
             Address     000b.fd13.9080
             Hello Time   2 sec  Max Age 20 sec  Forward Delay 15 sec
             Aging Time 300
Interface        Port ID                      Designated               Port ID
Name             Prio.Nbr      Cost Sts       Cost Bridge ID           Prio.Nbr
---------------- -------- --------- --- --------- -------------------- --------
Fa0/1            128.1           19 BLK        19 32769 000b.befa.eec0 128.1
Fa0/2            128.2           19 BLK        19 32769 000b.befa.eec0 128.2
Fa0/3            128.3           19 FWD         0 32769 0009.7c0b.e7c0 128.1
Fa0/4            128.4           19 BLK         0 32769 0009.7c0b.e7c0 128.2
Fa0/5            128.5           19 FWD        19 32769 000b.fd13.9080 128.5
Gi0/1            128.25           4 FWD        19 32769 000b.fd13.9080 128.25
```

STP Convergence: Summary

Cabrillo College

Recall that switches go through three steps for their initial convergence:

STP Convergence

Step 1 Elect one Root Bridge
Step 2 Elect Root Ports
Step 3 Elect Designated Ports

Also, all STP decisions are based on a the following predetermined sequence:

Five-Step decision Sequence

Step 1 - Lowest BID

Step 2 - Lowest Path Cost to Root Bridge

Step 3 - Lowest Sender BID

Step 4 – Lowest Port Priority

Step 5 - Lowest Port ID

STP Convergence: Summary

Cabrillo College

Example:

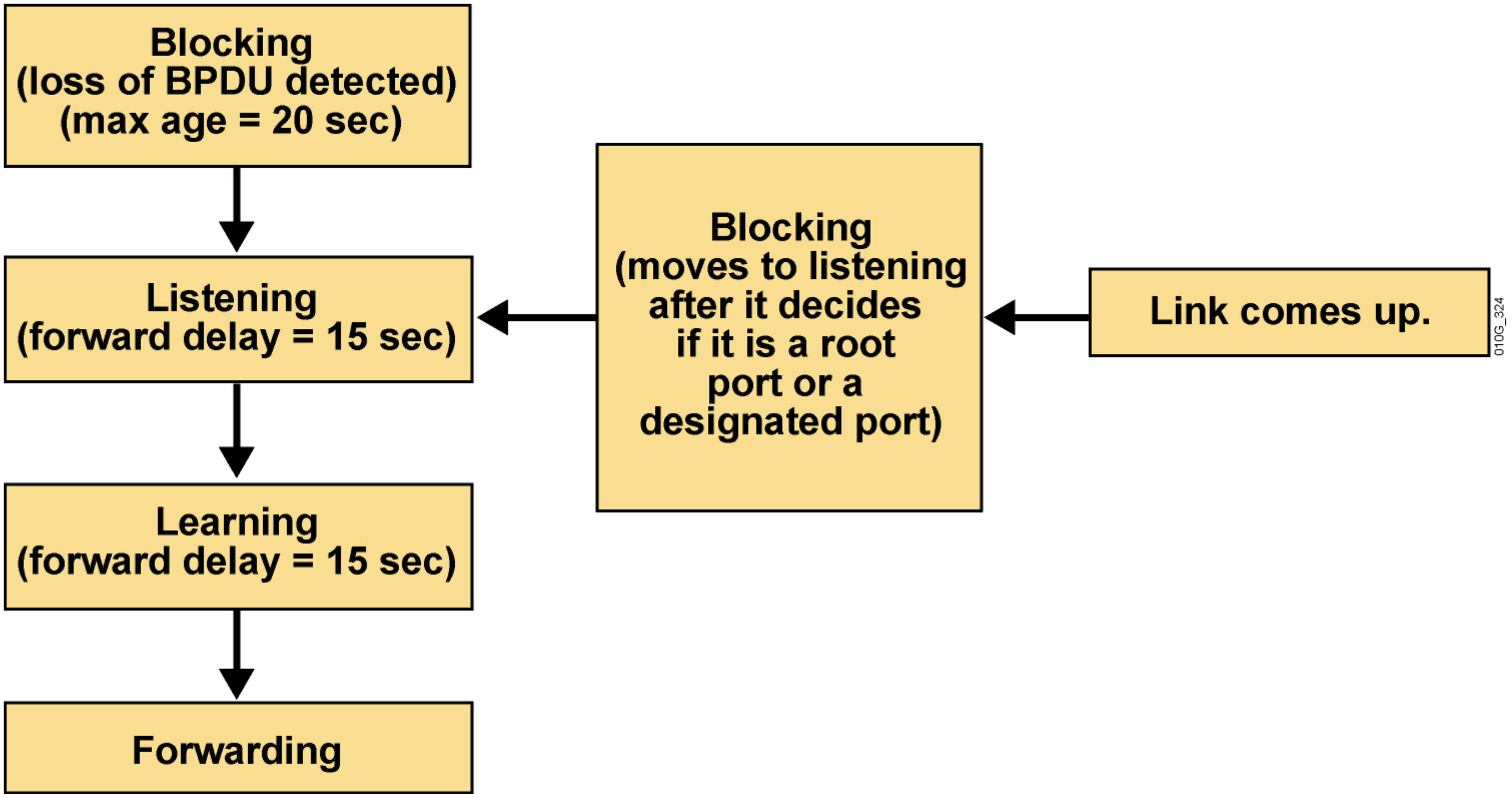
- A network that contains 15 switches and 146 segments (every switchport is a unique segment) would result in:
  - 1 Root Bridge
  - 14 Root Ports
  - 146 Designated Ports

Spanning-Tree Port States

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

Spanning Tree Port States

Spanning tree transitions each port through several different states.



Spanning-Tree Port States

Blocked:

- All ports start in blocked mode in order to prevent the bridge from creating a bridging loop.
- Port are listening (receiving) BPDUs.
- No user data is being passed.
- The port stays in a blocked state if Spanning Tree determines that there is a better path to the root bridge.
- May take a port up to 20 seconds to transition out of this state (max age). - coming soon.

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

Spanning-Tree Port States

Listen:

- The port transitions from the blocked state to the listen state
- Attempts to learn whether there are any other paths to the root bridge
- Listens to frames
- Port is not sending or receive user data
- Listens for a period of time called the forward delay (default 15 seconds).
- Ports that lose the *Designated Port election* become **non-Designated Ports** and drop back to **Blocking** state.

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

Spanning-Tree Port States

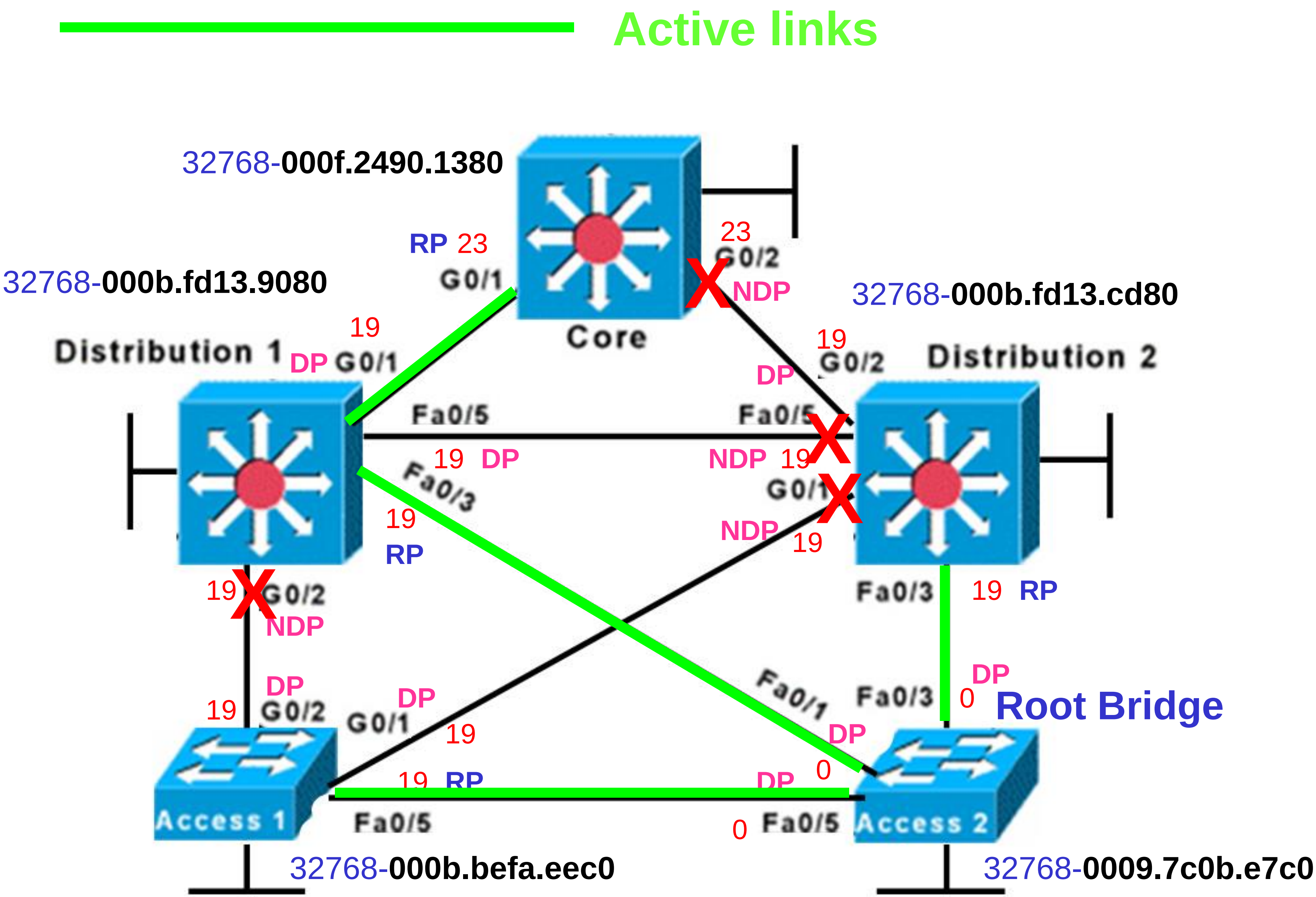
Cabrillo College

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

Designated Ports & Root Ports

Non-Designated Ports

Spanning-Tree Port States



Spanning-Tree Port States

Learn:

- The learn state is very similar to the listen state, except that the port can add information it has learned to its address table.
- Adds addresses to MAC Address Table
- Still not allowed to send or receive user data
- Learns for a period of time called the forward delay (default 15 seconds)

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

Spanning-Tree Port States

Forward:

- The port can send and receive user data.
- A port is placed in the forwarding state if:
  - There are no redundant links

  or

  - It is determined that it has the best path to the root

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

Spanning-Tree Port States

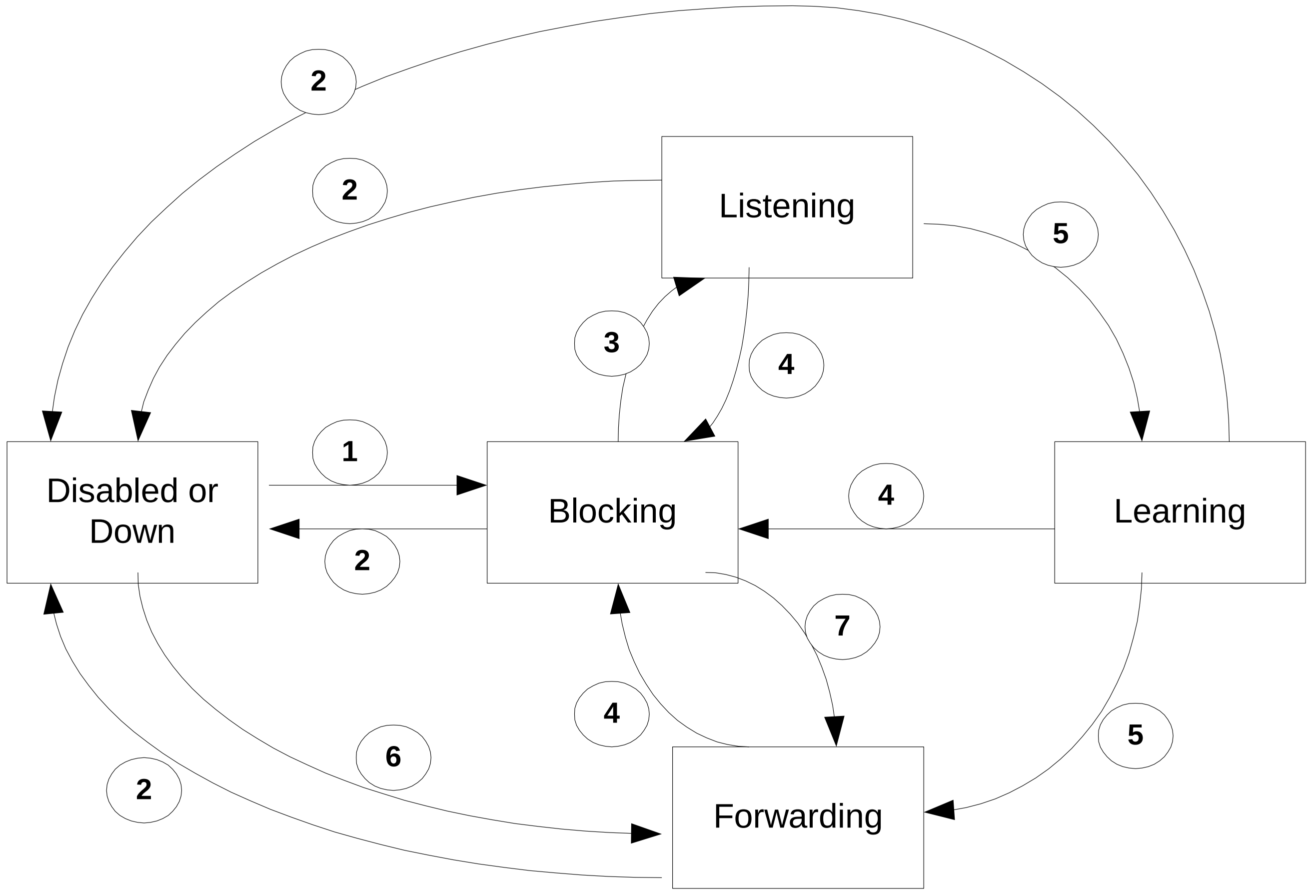
- Disabled: The port is shutdown.

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down

Results of BPDU exchange

- A **root port** for each switch and a designated port for each segment is selected.
  – These ports provide the best path from the switch to the root switch (usually the lowest-cost path).
  – These ports are put in the **forwarding** mode.
- Ports that will not be forwarding are placed in the **blocked** state.
  – These ports will continue to send and receive BPDU information but will not be allowed to send or receive data.

State	Purpose
Forwarding	Sending / receiving user data
Learning	Building bridging table
Listening	Building "active" topology
Blocking	Receives BPDUs only
Disabled	Administratively down



Standard States
(1) Port enabled or initialized
(2) Port disabled or failed
(3) Port selected as Root or Designated Port
(4) Port ceases to be a Root or Designated Port
(5) Forwarding timer expires

Cisco Specific States
(6) PortFast
(7) Uplink Fast

STP Timers

Timer	Primary Purpose	Default
Hello Time	Time between sending of Configuration BPDUs by the Root Bridge	2 Secs
Forward Delay	Duration of Listening and Learning States	15 Secs
Max Age	Time BPDU stored	20 Secs

STP Timers

Forward Delay Timer

- The default value of the forward delay (**15 seconds**) was originally derived assuming a maximum network size of seven bridge hops, a maximum of three lost BPDUs, and a hello-time interval of 2 seconds.
- See LAN Switching, by Clark, or other resources for this calculation
- Forward delay is used to determine the length of:
  - **Listening state**
  - **Learning state**

Timer	Primary Purpose	Default
Hello Time	Time between sending of Configuration BPDUs by the Root Bridge	2 Secs
Forward Delay	Duration of Listening and Learning States	15 Secs
Max Age	Time BPDU stored	20 Secs

STP Timers

Max Age Timer

- Max Age is the time that a bridge stores a BPDU before discarding it.
- Each port saves a copy of the best BPDU it has seen.
- If the device sending this best BPDU fails, it may take **20 seconds** the a switch transitions the connected port to Listening.

Timer	Primary Purpose	Default
Hello Time	Time between sending of Configuration BPDUs by the Root Bridge	2 Secs
Forward Delay	Duration of Listening and Learning States	15 Secs
Max Age	Time BPDU stored	20 Secs

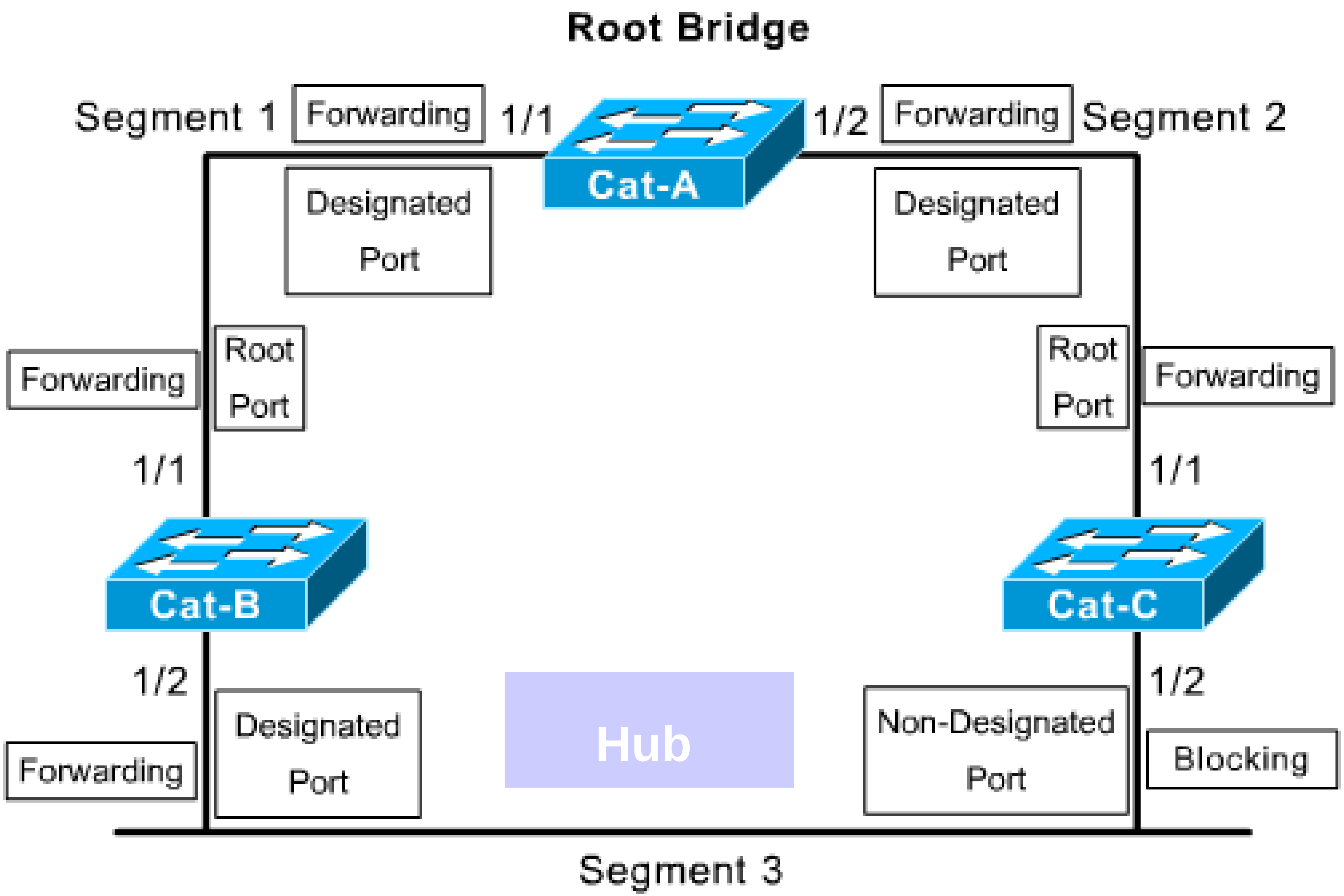
STP Timers

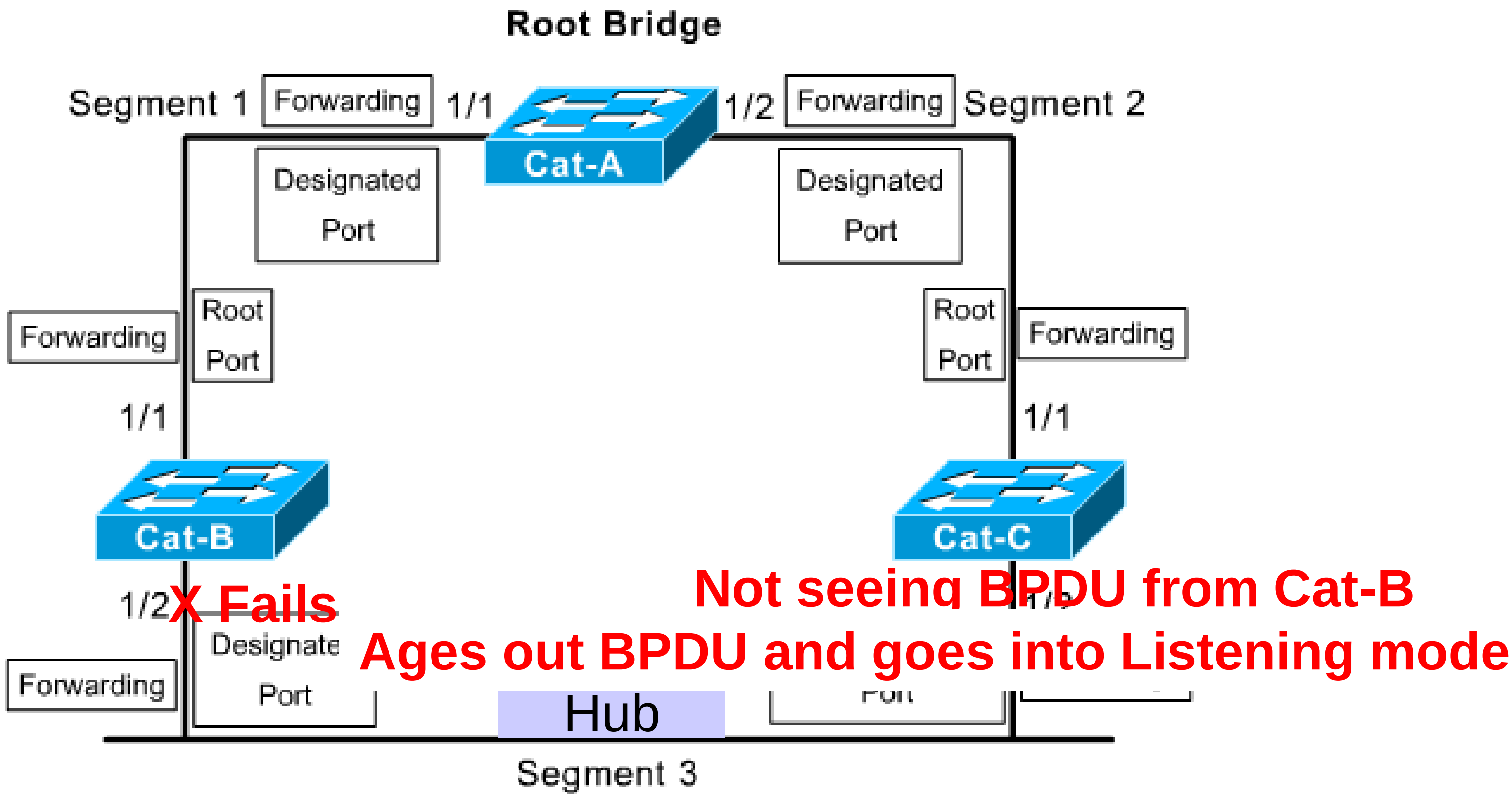
Modifying Timers

- Do not change the default timer values without careful consideration.
- Modify the STP timers only from the root bridge
- The BPDUs contain three fields where the timer values can be passed from the root bridge to all other bridges in the network.
- It can take 30-50 seconds for a switch to adjust to a change in topology.

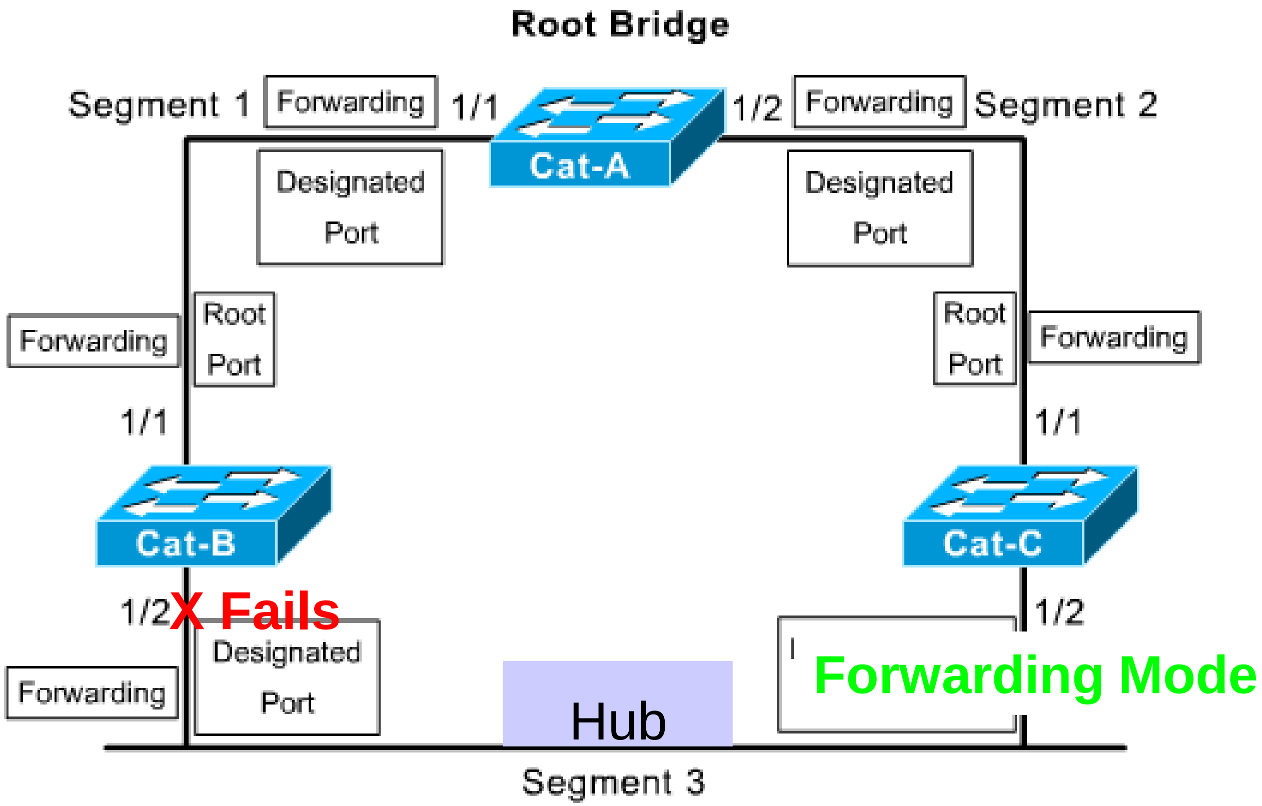
Timer	Primary Purpose	Default
Hello Time	Time between sending of Configuration BPDUs by the Root Bridge	2 Secs
Forward Delay	Duration of Listening and Learning States	15 Secs
Max Age	Time BPDU stored	20 Secs

Example


Root Bridge
Segment 1
Forwarding
1/1
Cat-A
1/2
Forwarding
Segment 2
Designated Port
Designated Port
Forwarding
Root Port
Root Port
Forwarding
1/1
1/1
Cat-B
Cat-C
1/2
1/2
Forwarding
Designated Port
Hub
Non-Designated Port
Blocking
Segment 3



- Cat-B:1/2 fails.
- Cat-C has no immediate notification because it's still receiving a link from the hub.
- Cat-C notices it is not receiving BPDUs from Cat-B.
- 20 seconds (max age) after the failure, Cat-C ages out the BPDU that lists Cat-B as having the DP for segment 3.
- This causes Cat-C:1/2 to transition into the Listing state in an effort to become the DP.



- Because Cat-C:1/2 now offers the most attractive access from the Root Bridge to this link, it eventually transitions all the way into Forwarding mode.
- In practice this will take 50 seconds (20 max age + 15 Listening + 15 Learning) for Cat-C:1/2 to take over after the failure of Cat-B:1/2.

BPDUs

Cabrillo College

- Configuration BPDUs
- Topology Change Notification (TCN) BPDUs

Configuration BPDUs

- A closer look...

Bytes	Field
2	Protocol ID
1	Version
1	Message type
1	Flags
8	Root ID
4	Cost of path
8	Bridge ID
2	Port ID
2	Message age
2	Max age
2	Hello time
2	Forward delay

310P_126

Configuration BPDUs

- **Protocol Identifier** (2 bytes), always 0
- **Version** (1 byte), always 0
- **Message Type** (1 byte): Determines whenter this is a Configuration BPDU or TCN BPDU
- **Flags** (1 byte): Used with topology changes. Used with TCN BPDUs
- **Root BID** (8 bytes): Indicates current Root Bridge on the network, includes:
  – *Bridge Priority* (2 bytes)
  – *Bridge MAC Address* (6 bytes)

Bytes	Field
2	Protocol ID
1	Version
1	Message type
1	Flags
8	Root ID
4	Cost of path
8	Bridge ID
2	Port ID
2	Message age
2	Max age
2	Hello time
2	Forward delay

310P_126

Configuration BPDUs

- **Root Path Cost (Cost to Root)** (4 bytes):
  - Cumulative cost of the path from the bridge sending the BDPU to the Root Bridge indicated in the Root ID field.
  - Cost is based on bandwidth.
- **(Sender's) Bridge ID** (8 bytes): Bridge ID sending the BDPU
  - 2 bytes: Bridge Priority
  - 6 bytes: MAC Address
- **Port ID** (2 bytes): Port on bridge sending BDPU, including Port Priority value. (0x800n - not necessarily consecutive.)

Bytes	Field
2	Protocol ID
1	Version
1	Message type
1	Flags
8	Root ID
4	Cost of path
8	Bridge ID
2	Port ID
2	Message age
2	Max age
2	Hello time
2	Forward delay

310P_126

Configuration BPDUs

Cabrillo College

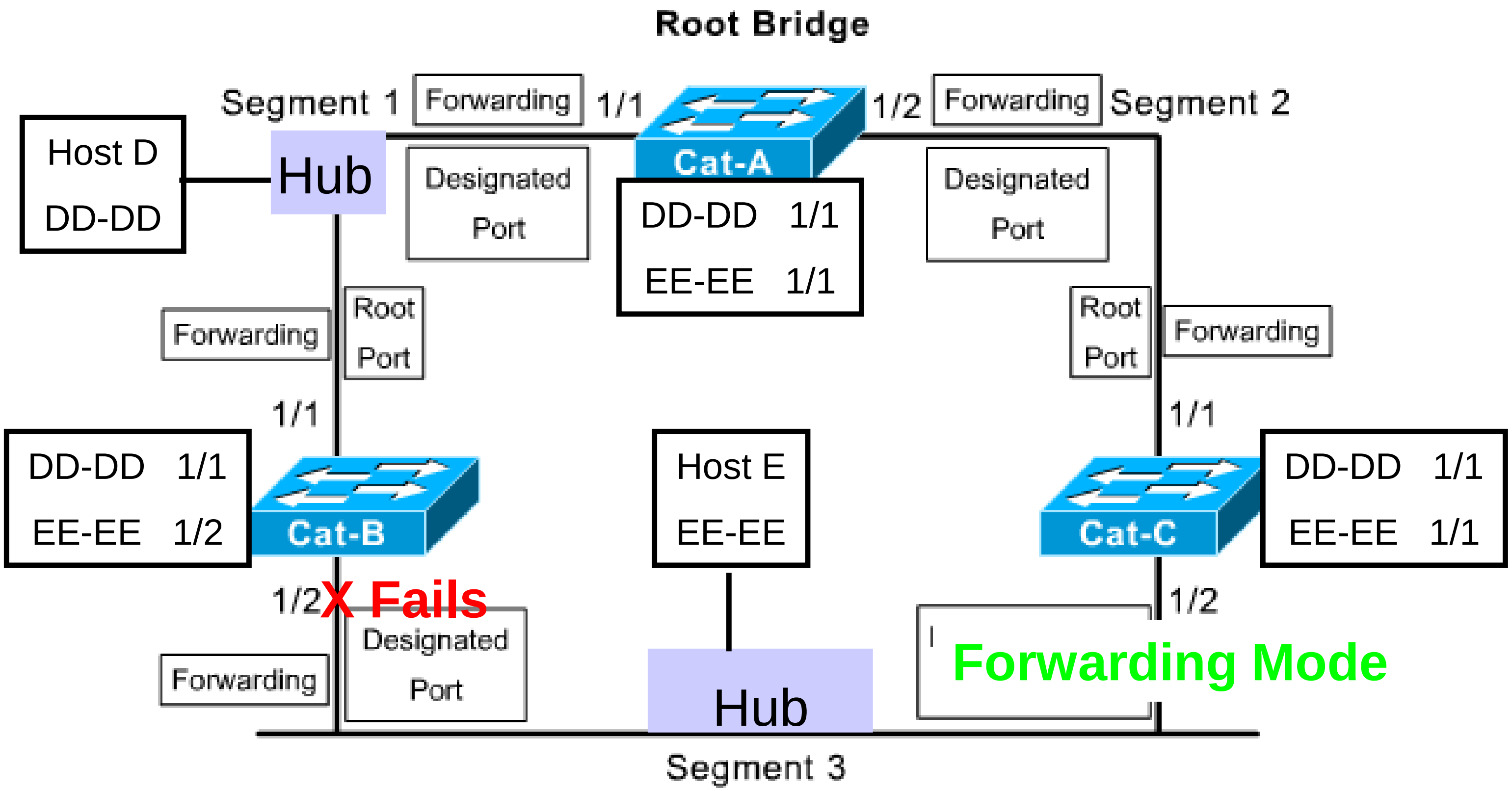
- **Message Age** (2 bytes):
  - Age of BDPU, encoded in 256ths of a second.
- **Maximum Age** (2 bytes):
  - When BDPU should be discarded (default 20 sec)
- **Hello Time** (2 bytes):
  - How often BDPU's are to be sent (default 2 sec)
- **Forward Delay** (2 bytes):
  - How long bridge should remain in listening and learning states (default 15 sec)

Bytes	Field
2	Protocol ID
1	Version
1	Message type
1	Flags
8	Root ID
4	Cost of path
8	Bridge ID
2	Port ID
2	Message age
2	Max age
2	Hello time
2	Forward delay

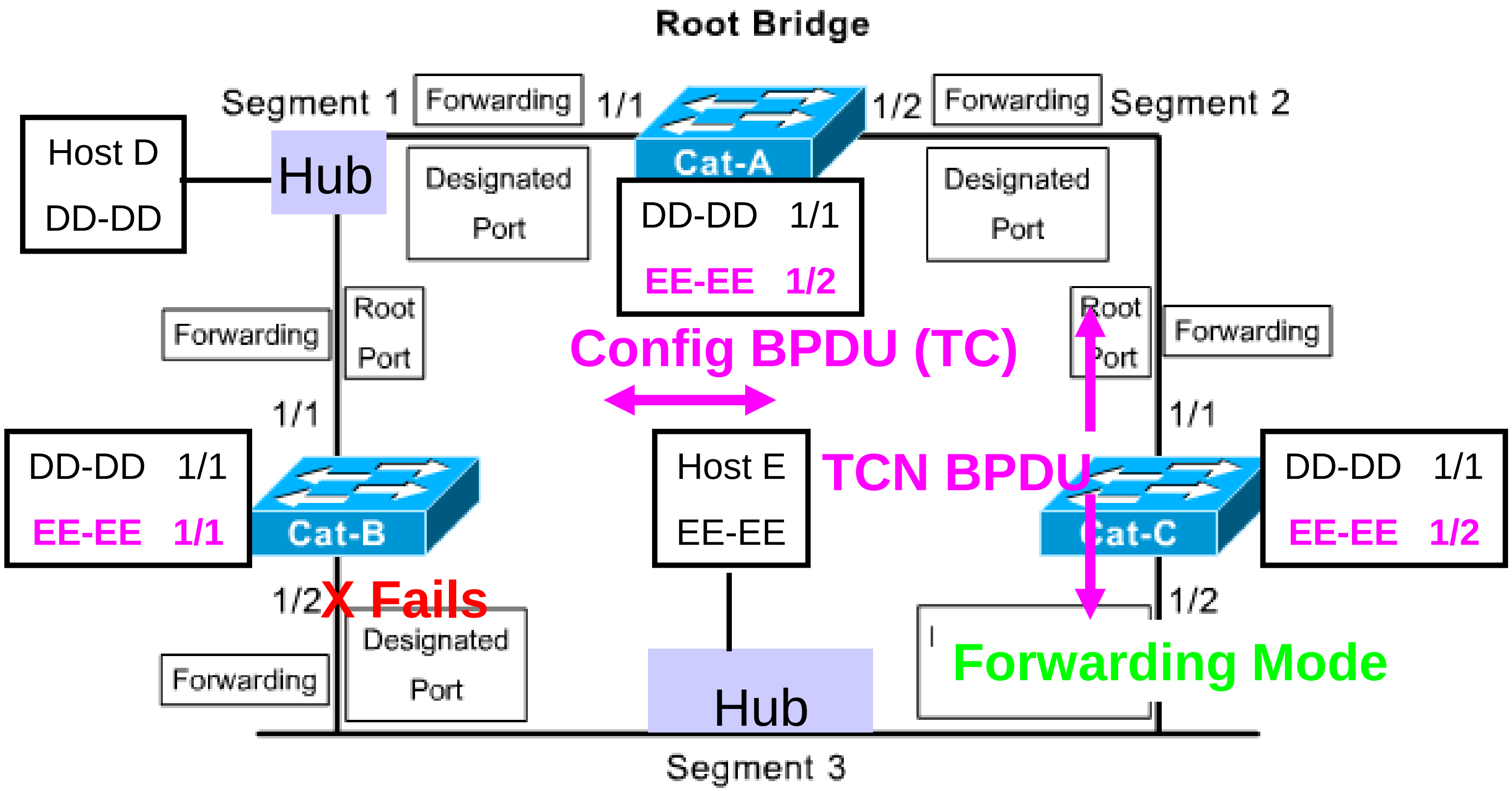
310P_126

TCN BPDUs

- It can take 30-50 seconds for a single switch to adjust to a change in topology.
- While the network is converging, physical addresses that can no longer be reached are still listed in the MAC address table.
- Because these addresses are in the table, the switch will attempt to forward frames to devices it cannot reach.



- Host-D is communicating with Host-E, via Cat-B, while Cat-B:1/2 fails.
- As discussed earlier, Cat-C:1/2 takes over as the DP in 50 seconds.
- However, without TCN BPDUs, the data traffic continues to be be sent to Cat-B for another 4 minutes and 10 seconds.
- Why? Prior to the failure notice the MAC Address Tables.
- All three switches have the traffic traveling counter-clockwise, because Cat-C:1/2 was in blocking mode.
- Although the change in Cat-C:1/2 is in forwarding mode, the MAC Address Tables are not correctly reflecting the change in STP topology.



- One option is to wait for the normal timeout of this entry in the MAC Address Table, which is 300 seconds (5 minutes). (This is where we got the 4 minutes and 10 seconds, plus 50 seconds for the STP timers.)
- A better solution is for switches to send out TCN BPDUs when there is a change in the forwarding state of a port, so switches age out their MAC Address Tables from 300 seconds to 15 seconds (Forward Delay).
- Doesn’t flush MAC Address Table, just accelerates the aging process.
- Devices that continue to speak for that 15 seconds will remain in the table.
- All other frames are flooded until the switch learns otherwise.

STP Topology Changes

- The STP change process requires the switch to clear the table faster in order to get rid of unreachable physical addresses.
- If a switch detects a change, it can send a Topology Change Notification (TCN) BPDU out its root port.
  - The topology change BPDU is **forwarded to the root switch**, and from there, is propagated throughout the network.
- **TCN does not start a STP recalculation.**

TCN BPDUs

Cabrillo College

- **Understanding Spanning-Tree Protocol Topology Changes** http://www.cisco.com/warp/public/473/17.html
- Remember that a TCN does not start a STP recalculation. This fear comes from the fact that TCNs are often associated with unstable STP environments; TCNs are a consequence of this, not a cause. The TCN only has an impact on the aging time; it will not change the topology nor create a loop.
- The number or the rate of topology changes is not an issue in itself. The problem is to know what the topology change means. A healthy network can experience a high rate of topology change. Nevertheless, ideally, a topology change would be related to a significant event in the network like a server going up or down or a link transitioning. This can be achieved by enabling portfast on ports that are going up and down as part of their normal operation.

STP Topology Changes

A Bridge originates a TCN BPDU in two conditions:

1. It transitions a port into Forwarding state and it has at lease one Designated Port or Root Port.
2. It transitions a port from either Forwarding or Learning states to the Blocking state.

- On bridges with Designated Ports accept and process TCN BPDUs.
- The Root Bridge will send out Configuration BPDUs
- See Clark, LAN Switching for more info.

- Much simpler than a Configuration BPDU.
- Only three fields, Protocol ID, Version, and Type (TCN).

Configuring the Root Bridge


```
Switch(config)#spanning-tree vlan 1 root primary
```

- **This command forces this switch to be the root.**

```
Switch(config)#spanning-tree vlan 1 root secondary
```

- **This command configures this switch to be the secondary root.**

Or

```
Switch(config)#spanning-tree vlan 1 priority priority
```

- **This command statically configures the priority (in increments of 4096).**

STP – Spanning Tree Protocol

Cabrillo College

CIS 187 Multilayer Switched Networks

CCNP 3 version 4

Rick Graziani

Fall 2006